GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 28, 2004, 09:07:19 ; Search time 126 Seconds
(without alignments)
953.037 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2219	100.0	425	4	AAU00438	Aau00438 Human neu
2	2218	100.0	425	4	AAB67489	Aab67489 Amino aci
3	2215	99.8	425	4	AAU11188	Aau11188 Human G p
4	2214	99.8	425	2	AAW80456	Aaw80456 G-protein
5	2214	99.8	425	4	AAU11186	Aau11186 Human G p
6	2214	99.8	425	4	AAB67079	Aab67079 Human HFG
7	2214	99.8	425	5	AAG78345	Aag78345 Human HFG
8	2214	99.8	425	6	ABP81941	Abp81941 Human ore
9	2209	99.5	425	4	ABB56378	Abb56378 Non-endog

10	2183	98.4	425	4	AAE04740	Aae04740 Cynomolgo
11	2097	94.5	402	2	AAW06124	Aaw06124 Neuropept
12	2093.5	94.3	401	5	AAG78346	Aag78346 Human HFG
13	2087	94.1	427	4	AAB47300	Aab47300 Dog orexi
14	1908	86.0	364	4	AAU00442	Aau00442 Human neu
15	1902.5	85.7	389	2	AAW80805	Aaw80805 Amino aci
16	1902.5	85.7	389	4	AAU11187	Aau11187 Human G p
17	1902.5	85.7	389	5	ABB08208	Abb08208 G-protein
18	1898	85.5	377	2	AAW06126	Aaw06126 Neuropept
19	1897	85.5	369	2	AAW06125	Aaw06125 Neuropept
20	1897	85.5	372	4	AAU00440	Aau00440 Human neu
21	1869	84.2	369	4	AAU00439	Aau00439 Human neu
22	1474.5	66.4	460	4	AAB61970	Aab61970 Rat HCRTR
23	1474.5	66.4	460	6	ABG73515	Abg73515 Rat OX2R
24	1469	66.2	444	4	AAB61968	Aab61968 Canine wi
25	1463	65.9	444	4	AAB84416	Aab84416 Amino aci
26	1460	65.8	444	4	AAB98007	Aab98007 Human hyp
27	1460	65.8	444	4	AAB61969	Aab61969 Human HCR
28	1460	65.8	444	6	ABG73514	Abg73514 Human OX2
29	1460	65.8	444	6	ABP81942	Abp81942 Human ore
30	1459	65.8	444	2	AAY03649	Aay03649 Human 7-t
31	1455	65.6	444	4	ABB56379	Abb56379 Non-endog
32	1311	59.1	263	2	AAR91233	Aar91233 Rabbit G-
33	1311	59.1	263	2	AAW11236	Aaw11236 G-protein
34	1164.5	52.5	330	4	AAB61971	Aab61971 Canine na
35	997	44.9	327	4	AAB61972	Aab61972 Canine na
36	522.5	23.5	430	3	AAY79376	Aay79376 Human neu
37	522.5	23.5	430	3	AAY93151	Aay93151 Novel hum
38	522.5	23.5	430	4	AAG67757	Aag67757 Amino aci
39	522.5	23.5	430	6	ABP97053	Abp97053 Human RFR
40	522.5	23.5	430	6	ABP81742	Abp81742 Human neu
41	522.5	23.5	430	7	ADB76177	Adb76177 Human OT7
42	522.5	23.5	430	7	ADD95536	Add95536 Human NPF
43	522.5	23.5	441	7	ADC86437	Adc86437 Human GPC
44	519.5	23.4	428	3	AAY56887	Aay56887 Human B5
45	519.5	23.4	428	6	ABG72069	Abg72069 Human G p

ALIGNMENTS

RESULT 1
AAU00438
ID AAU00438 standard; protein; 425 AA.
XX
AC AAU00438;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human neuropeptide receptor.
XX
KW Human; neuropeptide receptor; neuropeptide Y receptor; obesity;
KW nervous system disorder; hyperproliferative disorder; diabetes mellitus;
KW cardiovascular disorder; autoimmune disorder; infectious disorder;
KW eating behaviour disorder; narcolepsy; neurological disease;
KW narcotics addiction; nicotine addiction; alcohol addiction; gene therapy;
KW protein co-ordinate data.

XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 47. .72
FT /label= TM1
FT /note= "Transmembrane region 1"
FT Region 83. .106
FT /label= TM2
FT /note= "Transmembrane region 2"
FT Region 112. .142
FT /label= TM3
FT /note= "Transmembrane region 3"
FT Region 163. .189
FT /label= TM4
FT /note= "Transmembrane region 4"
FT Region 214. .239
FT /label= TM5
FT /note= "Transmembrane region 5"
FT Region 299. .327
FT /label= TM6
FT /note= "Transmembrane region 6"
FT Region 335. .363
FT /label= TM7
FT /note= "Transmembrane region 7"
XX
PN WO200117532-A1.
XX
PD 15-MAR-2001.
XX
PF 07-SEP-2000; 2000WO-US024518.
XX
PR 10-SEP-1999; 99US-00393696.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Soppet DR, Li Y, Rosen CA;
XX
DR WPI; 2001-183276/18.
DR N-PSDB; AAS00491.
XX
PT A new nucleic acid encoding a human neuropeptide receptor polypeptide,
PT useful for preventing, treating or ameliorating obesity, narcolepsy,
PT neurological disease and addiction to narcotics, nicotine and alcohol.
XX
PS Claim 3; Fig 4; 385pp; English.
XX
CC The present sequence represents a novel human neuropeptide receptor which
CC shows sequence homology to the neuropeptide Y receptor. Two splice
CC variants of the neuropeptide receptor (AAU00439-AAU00440) and a possible
CC mutant (AAU00442) are also described. Polypeptides and polynucleotides of
CC the neuropeptide receptor are useful for diagnosing, preventing, or
CC treating a pathological condition in a subject related to the central
CC nervous and peripheral nervous systems (CNS and PNS). The polypeptides
CC and polynucleotides may be used to treat hyperproliferative,
CC cardiovascular, autoimmune, nervous system or infectious disorders e.g.
CC cancer, heart disease, rheumatoid arthritis, Alzheimer's disease, HIV

CC infection and diabetes mellitus. In particular they are useful for
CC preventing, treating or ameliorating a medical condition in a mammal such
CC as obesity/eating behaviour disorders, narcolepsy, neurological disease,
CC addiction to narcotics, nicotine and alcohol, chronic pain, acute pain,
CC migraine headaches and anxiety disorders. The polynucleotides encoding
CC the neuropeptide receptor can also be used in gene therapy methods for
CC treating such diseases
XX
SQ Sequence 425 AA;

Query Match 100.0%; Score 2219; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 6.8e-221;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 2
AAB67489
ID AAB67489 standard; protein; 425 AA.
XX
AC AAB67489;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human hypocretin (orexin) receptor 1.
XX

KW Human; hypocretin receptor 1; orexin receptor 1; HCRTR1; chromosome 1;
KW 1p33; central nervous system modulator.
XX
OS Homo sapiens.
XX
PN WO200114555-A1.
XX
PD 01-MAR-2001.
XX
PF 22-AUG-2000; 2000WO-US022986.
XX
PR 23-AUG-1999; 99US-00379083.
PR 07-JAN-2000; 2000US-00479128.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Olafsdottir BR, Gulcher J;
XX
DR WPI; 2001-211306/21.
DR N-PSDB; AAF55159.
XX
PT Novel isolated nucleic acid molecule encoding hypocretin (orexin)
PT receptor 1 useful for treating and diagnosing narcolepsy.
XX
PS Disclosure; Fig 1A; 44pp; English.
XX
CC The present sequence represents a human hypocretin (orexin) receptor 1
CC (HCRTR1) polypeptide. The HCRTR1 gene is present on chromosome 1,
CC location 1p33. It is likely that a mutation in the HCRTR1 gene is
CC associated with narcolepsy. HCRTR1 is a central nervous system modulator.
CC The HCRTR1 polypeptide and polynucleotide are useful for diagnosing or
CC treating narcolepsy in an individual. The HCRTR1 polynucleotide is a
CC source of probes and primers, and is also used to produce the protein
CC recombinantly
XX
SQ Sequence 425 AA;

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 8.6e-221;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
```

```
Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 3
AAU11188
ID AAU11188 standard; protein; 425 AA.
XX
AC AAU11188;
XX
DT 25-FEB-2002 (first entry)
XX
DE Human G protein-coupled receptor HFGAN72X variant.
XX
KW Human; G protein-coupled receptor; GPCR; HFGAN72X; bacterial infection;
KW fungal infection; protozoan infection; viral infection;
KW human immunodeficiency virus; HIV; cancer; diabetes; Parkinson's disease;
KW osteoporosis; myocardial infarction; ulcer; asthma; allergy;
KW angina pectoris; renal disease; depression; schizophrenia; anorexia;
KW obesity; Kallman's syndrome; hypothalamic disorder;
KW idiopathic hormone deficiency; gigantism; migraine; pain; lung disease;
KW burn; sleep disorder; jet lag; Huntington's disease; gene therapy.
XX
OS Homo sapiens.
XX
PN US2001025031-A1.
XX
PD 27-SEP-2001.
XX
PF 06-APR-2001; 2001US-00828538.
XX
PR 08-JUN-1998; 98US-0088524P.
PR 22-JUL-1998; 98US-0093726P.
PR 08-JUN-1999; 99US-00328014.
XX
PA (ELLI/) ELLIS C E.
PA (KWOK/) KWOK C.
PA (BODS/) BODSWORTH N J.
PA (HALS/) HALSEY W.
PA (HORN/) HORN S V.
XX
PI Ellis CE, Kwok C, Bodsworth NJ, Halsey W, Horn SV;
XX
DR WPI; 2001-624968/72.

DR N-PSDB; AAS17464.
XX
PT Isolated HFGAN72 receptor useful for treatment of a patient having need
PT of HFGAN72 receptor and in the detection and treatment of disease, e.g.
PT infections such as bacterial, fungal, protozoan and viral infections and
PT cancers.
XX
PS Claim 23; Fig 6; 75pp; English.
XX
CC The invention relates to an isolated polypeptide, the HFGAN72 receptor or
CC its variant, encoded by the 8 exon sequences given in the specification.
CC HFGAN72 is a G protein-coupled receptor (GPCR). HFGAN72 is useful for the
CC treatment of a patient having need of HFGAN72 receptor where HFGAN72 is
CC administered by providing to the patient DNA encoding HFGAN72 and
CC expressing HFGAN72 in vivo (i.e by gene therapy). HFGAN72 is particularly
CC useful for applications in the detection and treatment of disease, e.g.
CC infections such as bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by human immunodeficiency virus (HIV)-1 or
CC HIV-2, cancers, diabetes, Parkinson's disease, osteoporosis, myocardial
CC infarction, ulcers, asthma, allergies, angina pectoris, renal disease,
CC depression, schizophrenia, anorexia, obesity, Kallman's syndrome,
CC hypothalamic disorders, idiopathic hormone deficiency (e.g. gigantism),
CC migraine, pain, lung diseases, burns, sleep disorders, jet lag,
CC Huntington's disease and many other diseases and disorders given in the
CC specification. The present sequence is the human HFGAN72X variant,
CC encoded by an alternative allele of HFGAN72
XX
SQ Sequence 425 AA;

Query Match 99.8%; Score 2215; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 1.8e-220;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
```

```
Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 4
AAW80456
ID AAW80456 standard; protein; 425 AA.
XX
AC AAW80456;
XX
DT 26-JAN-1999 (first entry)
XX
DE G-protein coupled receptor (HFGAN72X) polypeptide.
XX
KW G-protein coupled receptor; HFGAN72X; HIV infection; anorexia; cancer;
KW bulimia; asthma; Parkinson's disease; acute heart failure;
KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
KW benign prostatic hypertrophy; neurological disorder.
XX
OS Homo sapiens.
XX
PN EP875566-A2.
XX
PD 04-NOV-1998.
XX
PF 27-OCT-1997; 97EP-00308563.
XX
PR 30-APR-1997; 97US-00846704.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Bergsma DJ, Ellis CE;
XX
DR WPI; 1998-559432/48.
DR N-PSDB; AAV63468.
XX
PT New human G-protein coupled receptor HFGAN72X polypeptide and
PT polynucleotide - useful as diagnostic reagents and for treating e.g. HIV
PT infection, cancer and Parkinson's disease.
XX
PS Claim 1; Page 7-8; 24pp; English.
XX
CC The present sequence represents a G-protein coupled receptor (HFGAN72X)
CC polypeptide. HFGAN72X polypeptides and polynucleotides are useful for
CC diagnosing diseases related to over or under expression of HFGAN72X
CC proteins by identifying mutations in the HFGAN72X gene using HFGAN72X
CC probes, or determining HFGAN72X protein or mRNA expression levels.
CC HFGAN72X polypeptides are also useful for screening for compounds which
CC affect activity of the protein. Diseases that can be treated with
CC HFGAN72X include HIV infections, pain, anorexia, cancers, bulimia,
CC asthma, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, angina pectoris,

```
CC   myocardial infarction, ulcers, allergies, benign prostatic hypertrophy,
CC   and psychotic and neurological disorders
XX
SQ   Sequence 425 AA;

  Query Match             99.8%;  Score 2214;  DB 2;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 2.2e-220;
  Matches  423;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 5
AAU11186
ID   AAU11186 standard; protein; 425 AA.
XX
AC   AAU11186;
XX
DT   25-FEB-2002  (first entry)
XX
DE   Human G protein-coupled receptor HFGAN72X.
XX
KW   Human; G protein-coupled receptor; GPCR; HFGAN72X; bacterial infection;
KW   fungal infection; protozoan infection; viral infection;
KW   human immunodeficiency virus; HIV; cancer; diabetes; Parkinson's disease;
KW   osteoporosis; myocardial infarction; ulcer; asthma; allergy;
KW   angina pectoris; renal disease; depression; schizophrenia; anorexia;
```

KW obesity; Kallman's syndrome; hypothalamic disorder;
KW idiopathic hormone deficiency; gigantism; migraine; pain; lung disease;
KW burn; sleep disorder; jet lag; Huntington's disease; gene therapy.
XX
OS Homo sapiens.
XX
PN US2001025031-A1.
XX
PD 27-SEP-2001.
XX
PF 06-APR-2001; 2001US-00828538.
XX
PR 08-JUN-1998; 98US-0088524P.
PR 22-JUL-1998; 98US-0093726P.
PR 08-JUN-1999; 99US-00328014.
XX
PA (ELLI/) ELLIS C E.
PA (KWOK/) KWOK C.
PA (BODS/) BODSWORTH N J.
PA (HALS/) HALSEY W.
PA (HORN/) HORN S V.
XX
PI Ellis CE, Kwok C, Bodsworth NJ, Halsey W, Horn SV;
XX
DR WPI; 2001-624968/72.
XX
PT Isolated HFGAN72 receptor useful for treatment of a patient having need
PT of HFGAN72 receptor and in the detection and treatment of disease, e.g.
PT infections such as bacterial, fungal, protozoan and viral infections and
PT cancers.
XX
PS Claim 8; Fig 2; 75pp; English.
XX
CC The invention relates to an isolated polypeptide, the HFGAN72 receptor or
CC its variant, encoded by the 8 exon sequences given in the specification.
CC HFGAN72 is a G protein-coupled receptor (GPCR). HFGAN72 is useful for the
CC treatment of a patient having need of HFGAN72 receptor where HFGAN72 is
CC administered by providing to the patient DNA encoding HFGAN72 and
CC expressing HFGAN72 in vivo (i.e by gene therapy). HFGAN72 is particularly
CC useful for applications in the detection and treatment of disease, e.g.
CC infections such as bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by human immunodeficiency virus (HIV)-1 or
CC HIV-2, cancers, diabetes, Parkinson's disease, osteoporosis, myocardial
CC infarction, ulcers, asthma, allergies, angina pectoris, renal disease,
CC depression, schizophrenia, anorexia, obesity, Kallman's syndrome,
CC hypothalamic disorders, idiopathic hormone deficiency (e.g. gigantism),
CC migraine, pain, lung diseases, burns, sleep disorders, jet lag,
CC Huntington's disease and many other diseases and disorders given in the
CC specification. The present sequence is the human HFGAN72X receptor being
CC the product of a splice variant of HFGAN72
XX
SQ Sequence 425 AA;

Query Match 99.8%; Score 2214; DB 4; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.2e-220;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

```
RESULT 6
AAB67079
ID   AAB67079 standard; protein; 425 AA.
XX
AC   AAB67079;
XX
DT   10-APR-2001  (first entry)
XX
DE   Human HFGAN72 receptor SEQ ID NO: 13.
XX
KW   Human; mouse; rat; Lig72A; Lig72B; neuropeptide receptor; HFGAN72;
KW   truncation mutant; ligand; neurodegenerative disorder; pain;
KW   eating disorder; behaviour disorder; mood disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100787-A2.
XX
PD   04-JAN-2001.
XX
PF   22-JUN-2000; 2000WO-US017251.
XX
PR   25-JUN-1999;   99US-0141156P.
XX
```

PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Bingham S, Darker J, Liu W, Martin JD, Parsons AA, Patel SR;
XX
DR WPI; 2001-071483/08.
XX
PT Polynucleotides encoding Lig 72A polypeptides or their variants, which
PT are useful in the treatment of a disease or disorder associated with
PT pain, e.g. enhanced or exaggerated sensitivity to pain, hyperalgesia,
PT neuropathic pain and back pain.
XX
PS Claim 8; Fig 7; 101pp; English.
XX
CC The present invention provides the protein and coding sequences for the
CC human, mouse and rat HFGAN receptor ligand Lig72A. It also provides
CC truncated mutant versions. These, and their agonists and antagonists, are
CC all useful in the treatment of eating, neurodegenerative, behaviour,
CC mood, sexual, hormonal and sleep disorders, pain, depression, epilepsy
CC and acute inflammatory conditions
XX
SQ Sequence 425 AA;

Query Match 99.8%; Score 2214; DB 4; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.2e-220;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
```

Db 421 TTVLP 425

RESULT 7
AAG78345
ID AAG78345 standard; protein; 425 AA.
XX
AC AAG78345;
XX
DT 22-JAN-2002 (first entry)
XX
DE Human HFGAN72X G coupled receptor polypeptide.
XX
KW Antibacterial; fungicide; virucide; protozoacide; anti-HIV; analgesic;
KW cytostatic; nootropic; antiparkinsonian; cardiant; antiulcer;
KW antiasthmatic; tranquiliser; neuroleptic; antidepressant; anticonvulsant;
KW osteopathic; HIV infection; pain; cancer; anorexia; bulimia;
KW Parkinson's disease; acute heart failure; hypotension; hypertension;
KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
KW ulcers; asthma; allergy; delirium; dementia;
KW benign prostatic hypertrophy; anxiety; schizophrenia; manic depression;
KW dyskinesia; G coupled receptor; HFGAN72X; 7 transmembrane receptor.
XX
OS Homo sapiens.
XX
PN EP1154019-A2.
XX
PD 14-NOV-2001.
XX
PF 27-OCT-1997; 2001EP-00203008.
XX
PR 30-APR-1997; 97US-00846704.
PR 27-OCT-1997; 97EP-00308563.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Bergsma DJ, Ellis CE;
XX
DR WPI; 2002-012659/02.
DR N-PSDB; AAI64172.
XX
PT Nucleic acid encoding the HFGAN72X receptor, useful for diagnosis and
PT treatment of e.g. infections, cancer, anorexia, bulimia, Parkinson's
PT disease, and acute heart failure.
XX
PS Claim 11; Page 7-8; 24pp; English.
XX
CC The present sequence is that of a human HFGAN72X polypeptide encoded by a
CC cDNA shown in AAI64172. The specification describes a newly isolated
CC polynucleotide encoding a HFGAN72X G coupled receptor polypeptide. The
CC protein of the invention has antibacterial, fungicide, virucide,
CC protozoacide, anti-HIV, cardiant, analgesic, cytostatic, nootropic,
CC antiparkinsonian, antiulcer, antiasthmatic, tranquiliser, neuroleptic,
CC antidepressant, anticonvulsant and osteopathic activities. HFGAN72X
CC polynucleotides (PNs) are used to express HFGAN72X in vivo, to treat
CC diseases requiring increased activity or expression of HFGAN72X; for
CC recombinant production of HFGAN72X; diagnose diseases by detecting

CC mutations in genomic sequences and in chromosome identification and
CC mapping. HFGAN72X polypeptides are used to raise specific antibodies; as
CC therapeutic agents; to identify HFGAN72X protein-expressing clones; to
CC purify HFGAN72X proteins; in vaccines. Cells transformed with HFGAN72X
CC PNs are used to identify (ant)agonists of HFGAN72X, useful
CC therapeutically. Nucleic acids that inhibit expression of HFGAN72X and
CC polypeptides that compete with ligands for binding to HFGAN72X proteins
CC are also useful therapeutically and diagnostically. HFGAN72X-related
CC diseases include infections (bacterial, viral, fungal or protozoal,
CC particularly HIV-1 or -2); pain; cancer; anorexia; bulimia; Parkinson's
CC disease; acute heart failure; hypotension; hypertension; urinary
CC retention; osteoporosis; angina pectoris; myocardial infarction; ulcers;
CC asthma; allergy; benign prostatic hypertrophy; anxiety; schizophrenia;
CC manic depression; delirium; dementia; severe mental retardation and
CC dyskinesias
XX
SQ Sequence 425 AA;

Query Match 99.8%; Score 2214; DB 5; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.2e-220;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 8
ABP81941

ID ABP81941 standard; protein; 425 AA.
XX
AC ABP81941;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human orexin receptor 1 protein SEQ ID NO:368.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR N-PSDB; ABZ42789.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative

CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 425 AA;

Query Match 99.8%; Score 2214; DB 6; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.2e-220;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 9
ABB56378
ID ABB56378 standard; protein; 425 AA.
XX
AC ABB56378;

XX

DT 18-FEB-2002 (first entry)

XX

DE Non-endogenous human GPCR protein, SEQ ID NO: 549.

XX

KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KW constitutively activated GPCR; agonist; disease.

XX

OS Homo sapiens.
OS Synthetic.

XX

PN WO200177172-A2.

XX

PD 18-OCT-2001.

XX

PF 05-APR-2001; 2001WO-US011098.

XX

PR 07-APR-2000; 2000US-0195747P.

XX

PA (AREN-) ARENA PHARM INC.

XX

PI Lehmann-Bruinsma K, Liaw CW, Lin I;

XX

DR WPI; 2001-648759/74.
DR N-PSDB; ABI98014.

XX

PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with versions
PT of GPCRs.

XX

PS Claim 1; Page 350-351; 394pp; English.

XX

CC The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous constitutively
CC activated versions of known GPCRs are used in the invention for the
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists. Such agonists are useful as
CC therapeutic agents for diseases or disorders associated with GPCRs. The
CC present sequence is a non-endogenous version of a known human GPCR

XX

SQ Sequence 425 AA;

Query Match 99.5%; Score 2209; DB 4; Length 425;
Best Local Similarity 99.3%; Pred. No. 7.4e-220;
Matches 422; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
```

```
Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTKKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

```
RESULT 10
AAE04740
ID   AAE04740 standard; protein; 425 AA.
XX
AC   AAE04740;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Cynomolgous Monkey Orexin 1 Receptor.
XX
KW   Cynomolgous monkey; Orexin 1 Receptor; 7 Transmembrane Receptor family;
KW   7TM; gene therapy; vaccine; microbial infection; HIV-1; HIV-2; pain;
KW   cancer; diabetes; obesity; anorexia; bulimia; urinary retention;
KW   Parkinson's disease; acute heart failure; hypotension; hypertension;
KW   osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer;
KW   asthma; allergy; benign prostatic hypertrophy; migraine; vomiting;
KW   psychotic disorder; neurological disorder; anxiety; schizophrenia;
KW   manic depression; depression; delirium; dementia; mental retardation;
KW   dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
XX
OS   Macaca fascicularis.
XX
PN   WO200140259-A2.
XX
PD   07-JUN-2001.
XX
PF   04-DEC-2000; 2000WO-US032849.
XX
PR   02-DEC-1999;   99US-0168553P.
PR   28-NOV-2000; 2000US-00723781.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Ellis CE;
```

XX
DR WPI; 2001-408276/43.
DR N-PSDB; AAD09335.
XX
PT Novel Cynomolgous Monkey Orexin 1 Receptor polypeptides, for treating
PT infections, pain, cancer, diabetes, obesity, asthma, schizophrenia,
PT hypertension, urinary retention, Parkinson's disease and stroke.
XX
PS Claim 2; Page 28; 33pp; English.
XX
CC The present sequence is Cynomolgous Monkey Orexin 1 Receptor which is
CC structurally related to members of 7 Transmembrane Receptor (7TM) family.
CC The Orexin 1 Receptor polypeptide and polynucleotide are useful for
CC treating bacterial, fungal, protozoan and viral infections, particularly
CC infections caused by HIV-1 or HIV-2, pain, cancer, diabetes, obesity,
CC anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, angina pectoris,
CC myocardial infarction, stroke, ulcers, asthma, allergies, benign
CC prostatic hypertrophy, migraine, vomiting, psychotic and neurological
CC disorders including anxiety, schizophrenia, manic depression, depression,
CC delirium, dementia and severe mental retardation, and dyskinesias, such
CC as Huntington's disease or Gilles de la Tourette's syndrome. The
CC polypeptide is also useful for structure-based design of its agonist,
CC antagonist or inhibitor. The polynucleotide is useful for chromosome
CC localisation studies and in gene therapy. The Orexin 1 Receptor
CC polypeptide and polynucleotide are also useful as vaccines
XX
SQ Sequence 425 AA;

Query Match 98.4%; Score 2183; DB 4; Length 425;
Best Local Similarity 98.1%; Pred. No. 3.7e-217;
Matches 417; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||| || |||||||||||||||||||||||||||||||||||||||:||
Db          1 MEPSATPGAQMRVPTGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFLVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPVSLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAVMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGQPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
```

```
Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||| ||||||||||||||||||||||||||||:|||||||||
Db        361 LSGKFREQFKAAFSCCLPGPGPCGSLKAPSPRSSASHKSLSLQSRCSVSKLSEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 11
AAW06124
ID AAW06124 standard; protein; 402 AA.
XX
AC AAW06124;
XX
DT 22-FEB-1997 (first entry)
XX
DE Neuropeptide receptor.
XX
KW Human; neuropeptide receptor; drug screening; receptor-agonist;
KW receptor-antagonist; anorectic; antitumour; anticholesterolemic;
KW neuroprotective; anticonvulsant; hypotensive; sedative; diagnostic;
KW gene therapy.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Domain	47. .72
FT		/note= "Transmembrane region-1"
FT	Domain	83. .106
FT		/note= "Transmembrane region-2"
FT	Domain	112. .142
FT		/note= "Transmembrane region-3"
FT	Domain	163. .189
FT		/note= "Transmembrane region-4"
FT	Domain	214. .239
FT		/note= "Transmembrane region-5"
FT	Domain	299. .327
FT		/note= "Transmembrane region-6"
FT	Domain	335. .363
FT		/note= "Transmembrane region-7"

XX
PN WO9634877-A1.
XX
PD 07-NOV-1996.
XX
PF 05-MAY-1995; 95WO-US005616.
XX
PR 05-MAY-1995; 95WO-US005616.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Soppet DR, Li Y, Rosen CA;
XX
DR WPI; 1996-506094/50.
DR N-PSDB; AAT42826.
XX

PT Human neuro-peptide receptor polypeptide(s) - used to identify
PT antagonists and agonists to such polypeptide(s), e.g. in the treatment of
PT obesity, Alzheimer's disease, epilepsy, etc.
XX
PS Claim 1; Page 49-50; 77pp; English.
XX
CC The sequence represents a human adult hypothalamus neuropeptide receptor,
CC structurally related to the G-protein-coupled receptor family. Splice
CC variants are given in AAW06125-26. The receptor contains 7 transmembrane
CC regions. The receptor may be produced in recombinant form and used in a
CC drug screening assay for isolation of receptor-agonists and -antagonists,
CC which may be used as anorectic, antitumour, anticholesterolemic,
CC neuroprotective, anticonvulsant, hypotensive or sedative drugs, etc. The
CC encoding DNA may be used in genetic disease diagnosis or gene therapy.
CC The receptor itself and its corresponding antibody may also be used in
CC therapy and diagnosis
XX
SQ Sequence 402 AA;

Query Match 94.5%; Score 2097; DB 2; Length 402;
Best Local Similarity 99.5%; Pred. No. 2.8e-208;
Matches 400; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCHERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
```

RESULT 12
AAG78346
ID AAG78346 standard; protein; 401 AA.
XX
AC AAG78346;

XX

DT 22-JAN-2002 (first entry)

XX

DE Human HFGAN72X G coupled receptor partial protein sequence.

XX

KW Antibacterial; fungicide; virucide; protozoacide; anti-HIV; analgesic;
KW cytostatic; nootropic; antiparkinsonian; cardiant; antiulcer;
KW antiasthmatic; tranquiliser; neuroleptic; antidepressant; anticonvulsant;
KW osteopathic; HIV infection; pain; cancer; anorexia; bulimia;
KW Parkinson's disease; acute heart failure; hypotension; hypertension;
KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
KW ulcers; asthma; allergy; delirium; dementia;
KW benign prostatic hypertrophy; anxiety; schizophrenia; manic depression;
KW dyskinesia; G coupled receptor; HFGAN72X; 7 transmembrane receptor.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers
FT Misc-difference 280
FT /note= "Encoded by GGC"

XX

PN EP1154019-A2.

XX

PD 14-NOV-2001.

XX

PF 27-OCT-1997; 2001EP-00203008.

XX

PR 30-APR-1997; 97US-00846704.
PR 27-OCT-1997; 97EP-00308563.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.

XX

PI Bergsma DJ, Ellis CE;

XX

DR WPI; 2002-012659/02.
DR N-PSDB; AAI64173.

XX

PT Nucleic acid encoding the HFGAN72X receptor, useful for diagnosis and
PT treatment of e.g. infections, cancer, anorexia, bulimia, Parkinson's
PT disease, and acute heart failure.

XX

PS Disclosure; Page 10; 24pp; English.

XX

CC The present sequence is that of a human HFGAN72X polypeptide encoded by a
CC cDNA shown in AAI64173. The specification describes a newly isolated
CC polynucleotide encoding a HFGAN72X G coupled receptor polypeptide. The
CC protein of the invention has antibacterial, fungicide, virucide,
CC protozoacide, anti-HIV, cardiant, analgesic, cytostatic, nootropic,
CC antiparkinsonian, antiulcer, antiasthmatic, tranquiliser, neuroleptic,
CC antidepressant, anticonvulsant and osteopathic activities. HFGAN72X
CC polynucleotides (PNs) are used to express HFGAN72X in vivo, to treat
CC diseases requiring increased activity or expression of HFGAN72X; for
CC recombinant production of HFGAN72X; diagnose diseases by detecting
CC mutations in genomic sequences and in chromosome identification and
CC mapping. HFGAN72X polypeptides are used to raise specific antibodies; as
CC therapeutic agents; to identify HFGAN72X protein-expressing clones; to
CC purify HFGAN72X proteins; in vaccines. Cells transformed with HFGAN72X

CC PNs are used to identify (ant)agonists of HFGAN72X, useful
CC therapeutically. Nucleic acids that inhibit expression of HFGAN72X and
CC polypeptides that compete with ligands for binding to HFGAN72X proteins
CC are also useful therapeutically and diagnostically. HFGAN72X-related
CC diseases include infections (bacterial, viral, fungal or protozoal,
CC particularly HIV-1 or -2); pain; cancer; anorexia; bulimia; Parkinson's
CC disease; acute heart failure; hypotension; hypertension; urinary
CC retention; osteoporosis; angina pectoris; myocardial infarction; ulcers;
CC asthma; allergy; benign prostatic hypertrophy; anxiety; schizophrenia;
CC manic depression; delirium; dementia; severe mental retardation and
CC dyskinesias
XX
SQ Sequence 401 AA;

Query Match 94.3%; Score 2093.5; DB 5; Length 401;
Best Local Similarity 99.8%; Pred. No. 6.5e-208;
Matches 401; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPR-RAFLAEVKQMRARRKTAKML 299

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 359

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
              |||||||||||||||||||||||||||||||||||||||||||
Db        360 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 401
```

RESULT 13
AAB47300
ID AAB47300 standard; protein; 427 AA.
XX
AC AAB47300;
XX
DT 22-AUG-2001 (first entry)
XX
DE Dog orexin 1 receptor.
XX

KW Beagle; dog; orexin 1 receptor; human immunodeficiency virus; HIV; pain;
KW cancer; diabetes; anorexia; bulimia; asthma; angina pectoris;
KW Parkinson's disease; acute heart failure; hypotension; schizophrenia;
KW hypertension; urinary retention; osteoporosis; Huntington's disease;
KW myocardial infarction; stroke; ulcer; allergy; depression; delirium;
KW benign prostatic hypertrophy; migraine; vomiting; psychosis; anxiety;
KW dementia; severe mental retardation; dyskinesia.
XX
OS Canis familiaris.
XX
PN WO200142268-A1.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US033106.
XX
PR 07-DEC-1999; 99US-0169373P.
PR 06-DEC-2000; 2000US-00169373.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Ellis CE;
XX
DR WPI; 2001-381627/40.
DR N-PSDB; AAC85974.
XX
PT Novel Beagle Orexin 1 receptor polypeptides and polynucleotides for
PT identifying modulators for use in preventing, and treating psychotic and
PT neurological disorders, asthma, cancer, diabetes and microbial
PT infections.
XX
PS Claim 1; Page 33; 36pp; English.
XX
CC This sequence represents Beagle Orexin 1 receptor. The orexin 1
CC polypeptide is useful in identifying compounds that may be agonists
CC and/or antagonists. The compounds identified are useful in preventing and
CC treating human diseases, including bacterial, fungal, protozoan and viral
CC infections, particularly, infections caused by human immunodeficiency
CC virus (HIV), pain, cancer, diabetes, anorexia, bulimia, asthma,
CC Parkinson's diseases, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC stroke, ulcer, allergies, benign prostatic hypertrophy, migraine,
CC vomiting, psychotic and neurological disorders, including anxiety,
CC schizophrenia, depression, delirium, dementia and severe mental
CC retardation and dyskinesias, such as Huntington's disease
XX
SQ Sequence 427 AA;

Query Match 94.1%; Score 2087; DB 4; Length 427;
Best Local Similarity 94.1%; Pred. No. 3.3e-207;
Matches 402; Conservative 5; Mismatches 18; Indels 2; Gaps 1;

```
Qy          1 MEPSATPGAQMGVPPGSRE--PSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFV 58
              |||||||||| | | |  |  || ||||||||||||||||||||||||||||||||||:
Db          1 MEPSATPGAQTGTPTGGGELSPSLVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFL 60
```

```
Qy          59 VALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHAL 118
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          61 VALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHTL 120

Qy         119 CKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVP 178
               |||||||||||||||||||||||||||||||||||||||| |||||||||||:|||
Db         121 CKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARSSILGIWAVSLAVMVP 180

Qy         179 QAAVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQI 238
               |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         181 QAAVMECSSVLPELANRTRLFSVCDEHWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQI 240

Qy         239 FRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAK 298
               ||||||||||||||||||||||||| |   ||| || ||||||||||||||||||||
Db         241 FRKLWGRQIPGTTSALVRNWKRPSDQLEDQGPGLSAEPPPRARAFLAEVKQMRARRKTAK 300

Qy         299 MLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIY 358
               ||||:|||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db         301 MLMVLLLVFALCYLPISVLNVLKRVFGMFRQSSDREAVYACFTFSHWLVYANSAANPIIY 360

Qy         359 NFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLT 418
               ||||||||||||||||||||||||| ||||||||||||||| |||||||: ||||||
Db         361 NFLSGKFREQFKAAFSCCLPGLGPCGSPKAPSPRSSASHKSLSLHSRCSVSKVPEHVVLT 420

Qy         419 SVTTVLP 425
               |||||||
Db         421 SVTTVLP 427
```

```
RESULT 14
AAU00442
ID   AAU00442 standard; protein; 364 AA.
XX
AC   AAU00442;
XX
DT   18-JUN-2001  (first entry)
XX
DE   Human neuropeptide receptor N-terminal and C-terminal deletion mutant.
XX
KW   Human; neuropeptide receptor; neuropeptide Y receptor; obesity;
KW   nervous system disorder; hyperproliferative disorder; diabetes mellitus;
KW   cardiovascular disorder; autoimmune disorder; infectious disorder;
KW   eating behaviour disorder; narcolepsy; neurological disease;
KW   narcotics addiction; nicotine addiction; alcohol addiction; gene therapy;
KW   protein co-ordinate data; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1. .364
FT                   /note= "Corresponds to wild type neuropeptide receptor
FT                   residues 17-380"
XX
PN   WO200117532-A1.
XX
PD   15-MAR-2001.
```

XX
PF 07-SEP-2000; 2000WO-US024518.
XX
PR 10-SEP-1999; 99US-00393696.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Soppet DR, Li Y, Rosen CA;
XX
DR WPI; 2001-183276/18.
XX
PT A new nucleic acid encoding a human neuropeptide receptor polypeptide,
PT useful for preventing, treating or ameliorating obesity, narcolepsy,
PT neurological disease and addiction to narcotics, nicotine and alcohol.
XX
PS Example 9; Page; 385pp; English.
XX
CC The present sequence represents a human neuropeptide receptor mutant
CC protein which is constructed by the deletion of N-terminal residues 1-16
CC and C-terminal residues 381-425 of the wild type novel neuropeptide
CC receptor protein. The novel neuropeptide receptor (AAU00438) shows
CC sequence homology to the neuropeptide Y receptor. Two splice variants of
CC the neuropeptide receptor (AAU00439-AAU00440) and a possible.
CC Polypeptides and polynucleotides of the neuropeptide receptor are useful
CC for diagnosing, preventing, or treating a pathological condition in a
CC subject related to the central nervous and peripheral nervous systems
CC (CNS and PNS). The polypeptides and polynucleotides may be used to treat
CC hyperproliferative, cardiovascular, autoimmune, nervous system or
CC infectious disorders e.g. cancer, heart disease, rheumatoid arthritis,
CC Alzheimer's disease, HIV infection and diabetes mellitus. In particular
CC they are useful for preventing, treating or ameliorating a medical
CC condition in a mammal such as obesity/eating behaviour disorders,
CC narcolepsy, neurological disease, addiction to narcotics, nicotine and
CC alcohol, chronic pain, acute pain, migraine headaches and anxiety
CC disorders. The polynucleotides encoding the neuropeptide receptor can
CC also be used in gene therapy methods for treating such diseases. Note the
CC present sequence is not given in the patent but is indexed from the wild
CC type sequence shown in Figure 4 (AAU00438)
XX
SQ Sequence 364 AA;

Query Match 86.0%; Score 1908; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 9.7e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 SREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHH 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHH 60

Qy         77 MRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLT 120

Qy        137 LSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRT 180
```

```
Qy        197 RLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVR 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVR 240

Qy        257 NWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISV 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISV 300

Qy        317 LNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC 360

Qy        377 LPGL 380
              ||||
Db        361 LPGL 364
```

RESULT 15
AAW80805
ID AAW80805 standard; protein; 389 AA.
XX
AC AAW80805;
XX
DT 29-JAN-1999 (first entry)
XX
DE Amino acid sequence of HFGAN72Y a G-protein coupled receptor.
XX
KW G-protein coupled receptor family; HFGAN72Y; mutation; probe; agonist;
KW antagonist; activation; inhibition; gene therapy; antibody;
KW immune response; vaccine; HIV-1; HIV-2; cancer; anorexia; bulimia;
KW asthma; Parkinson's disease; acute heart failure; hypotension;
KW hypertension; urinary retention; osteoporosis; angina pectoris;
KW myocardial infarction; ulcer; allergies; psychotic disorder;
KW neurological disorder; gene mapping.
XX
OS Homo sapiens.
XX
PN EP875565-A2.
XX
PD 04-NOV-1998.
XX
PF 27-OCT-1997; 97EP-00308554.
XX
PR 30-APR-1997; 97US-00846705.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Bergsma DJ, Ellis C;
XX
DR WPI; 1998-570286/49.
DR N-PSDB; AAV68511.
XX
PT New G-protein coupled receptor HFGAN72Y polypeptide and polynucleotide -
PT useful as diagnostic reagents and for prevention and treatment of HIV
PT infections, cancer, osteoporosis and Parkinson's disease.
XX
PS Claim 1; Page 7; 22pp; English.

XX
CC This is the amino acid sequence of the G-protein coupled receptor,
CC HFGAN72Y used in the method of the invention. HFGAN72Y polypeptides and
CC polynucleotides are useful for diagnosing susceptibility to diseases by
CC detecting mutations in the HFGAN72Y gene using probes containing the
CC HFGAN72Y nucleotide sequence, and can diagnose diseases associated with
CC HFGAN72Y imbalance by determining HFGAN72Y polypeptide or mRNA expression
CC levels. Agonists/antagonists can be used in treatment to activate/inhibit
CC HFGAN72Y activity, in addition to direct administration of antisense
CC sequences to prevent expression, or HFGAN72Y polypeptides to treat
CC conditions associated with a lack HFGAN72Y protein. Gene therapy may also
CC be used to affect endogenous HFGAN72Y polypeptide production. HFGAN72Y
CC antibodies are useful for inducing an immune response to immunise and
CC prevent diseases, and for isolating HFGAN72Y clones or purifying the
CC polypeptides by affinity chromatography. HFGAN72Y polypeptides can be
CC administered directly or as a vaccine to inoculate against diseases.
CC Diseases diagnosed, prevented or treated include HIV-1 or HIV-2
CC infections, pain, cancers, anorexia, bulimia, asthma, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, myocardial infarction, ulcers;
CC allergies, benign prostatic hypertrophy, and psychotic and neurological
CC disorders. The HFGAN72Y polypeptide is also useful for mapping the gene
CC to a chromosome, allowing gene inheritance to be studied through linkage
CC analysis
XX
SQ Sequence 389 AA;

Query Match 85.7%; Score 1902.5; DB 2; Length 389;
Best Local Similarity 96.1%; Pred. No. 4e-188;
Matches 368; Conservative 2; Mismatches 10; Indels 3; Gaps 2;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSC-CLPGLGP 382
              |||   ::   | ||   ||  |
```

Db 361 LSG--CKEKSLALSCPSCPGHDP 381

Search completed: September 28, 2004, 09:46:56
Job time : 130 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 28, 2004, 09:44:56 ; Search time 34 Seconds
(without alignments)
645.325 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2218	100.0	425	4	US-09-479-128-2	Sequence 2, Appli
2	2214	99.8	425	3	US-08-846-704-2	Sequence 2, Appli
3	2214	99.8	425	4	US-09-211-823C-22	Sequence 22, Appl
4	2108	95.0	402	3	US-08-846-704-4	Sequence 4, Appli
5	2104	94.8	402	4	US-08-462-509B-2	Sequence 2, Appli
6	2104	94.8	402	5	PCT-US95-05616-2	Sequence 2, Appli
7	1902.5	85.7	389	2	US-08-846-705-2	Sequence 2, Appli
8	1902.5	85.7	389	4	US-09-211-823C-23	Sequence 23, Appl
9	1901	85.7	377	5	PCT-US95-05616-6	Sequence 6, Appli
10	1897	85.5	369	4	US-08-462-509B-4	Sequence 4, Appli
11	1897	85.5	369	5	PCT-US95-05616-4	Sequence 4, Appli

12	1894	85.4	372	4	US-08-462-509B-6	Sequence 6, Appli
13	1460	65.8	444	4	US-09-426-290-2	Sequence 2, Appli
14	1459	65.8	444	3	US-09-119-788-2	Sequence 2, Appli
15	1311	59.1	263	3	US-08-513-974B-54	Sequence 54, Appl
16	1311	59.1	263	3	US-08-513-974B-376	Sequence 376, App
17	1311	59.1	263	4	US-09-461-436B-54	Sequence 54, Appl
18	522.5	23.5	430	3	US-09-255-368-8	Sequence 8, Appli
19	507.5	22.9	432	3	US-09-255-368-2	Sequence 2, Appli
20	499.5	22.5	420	3	US-09-255-368-6	Sequence 6, Appli
21	436	19.6	370	3	US-09-172-353-2	Sequence 2, Appli
22	436	19.6	370	3	US-09-172-353-3	Sequence 3, Appli
23	436	19.6	370	4	US-09-799-955-2	Sequence 2, Appli
24	436	19.6	370	4	US-09-799-955-3	Sequence 3, Appli
25	435.5	19.6	381	2	US-08-687-355A-4	Sequence 4, Appli
26	435.5	19.6	381	4	US-09-407-367-4	Sequence 4, Appli
27	434	19.6	370	3	US-08-513-974B-26	Sequence 26, Appl
28	434	19.6	370	3	US-08-513-974B-323	Sequence 323, App
29	434	19.6	370	3	US-09-172-353-5	Sequence 5, Appli
30	434	19.6	370	3	US-08-776-971-21	Sequence 21, Appl
31	434	19.6	370	3	US-08-776-971-104	Sequence 104, App
32	434	19.6	370	4	US-09-799-955-5	Sequence 5, Appli
33	434	19.6	370	4	US-09-461-436B-26	Sequence 26, Appl
34	434	19.6	381	1	US-08-192-288-2	Sequence 2, Appli
35	434	19.6	381	2	US-08-687-355A-2	Sequence 2, Appli
36	434	19.6	381	4	US-09-200-673-16	Sequence 16, Appl
37	434	19.6	381	4	US-09-407-367-2	Sequence 2, Appli
38	432.5	19.5	370	3	US-09-172-353-7	Sequence 7, Appli
39	432.5	19.5	370	4	US-09-799-955-7	Sequence 7, Appli
40	429.5	19.4	381	2	US-08-687-355A-6	Sequence 6, Appli
41	429.5	19.4	381	4	US-09-407-367-6	Sequence 6, Appli
42	428.5	19.3	428	1	US-08-570-157-5	Sequence 5, Appli
43	428.5	19.3	428	3	US-08-029-170-31	Sequence 31, Appl
44	428.5	19.3	428	4	US-09-076-510-5	Sequence 5, Appli
45	428.5	19.3	428	4	US-09-004-349-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-479-128-2
; Sequence 2, Application US/09479128
; Patent No. 6319710
; GENERAL INFORMATION:
;  APPLICANT:  Berglind Ran Olafsdottir
;  APPLICANT:  Jeffrey Gulcher
;  TITLE OF INVENTION:  HUMAN NARCOLEPSY GENE
;  FILE REFERENCE:  2345.1005-001
;  CURRENT APPLICATION NUMBER: US/09/479,128
;  CURRENT FILING DATE:  2000-01-07
;  PRIOR APPLICATION NUMBER:  US 09/379,083
;  PRIOR FILING DATE:  1999-08-23
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 425
;   TYPE: PRT
```

```
;   ORGANISM: Homo Sapiens
US-09-479-128-2

  Query Match             100.0%;  Score 2218;  DB 4;  Length 425;
  Best Local Similarity   99.8%;  Pred. No. 7.7e-193;
  Matches  424;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 2
US-08-846-704-2
; Sequence 2, Application US/08846704
; Patent No. 6020157
;  GENERAL INFORMATION:
;    APPLICANT:  BERGSMA, DERK J.
;    APPLICANT:  ELLIS, CATHERINE E.
;    TITLE OF INVENTION:  NOVEL G-PROTEIN COUPLED
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/846,704
;       FILING DATE:  30-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PRESTIA, PAUL F
;       REGISTRATION NUMBER:  23,031
;       REFERENCE/DOCKET NUMBER:  GH-70002
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610-407-0700
;       TELEFAX:  610-407-0701
;       TELEX:  846169
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  425 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-846-704-2

  Query Match             99.8%;  Score 2214;  DB 3;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 1.8e-192;
  Matches  423;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
                |||||
Db        421 TTVLP 425


RESULT 3
US-09-211-823C-22
; Sequence 22, Application US/09211823C
; Patent No. 6664229
; GENERAL INFORMATION:
;  APPLICANT:  HAGEN, JAMES JOSEPH
;  APPLICANT:  TERRETT, JONATHAN ALEXANDER
;  APPLICANT:  UPTON, NEIL
;  APPLICANT:  PIPER, DAVID
;  APPLICANT:  SMITH, MARTIN IAN
;  APPLICANT:  KENNETT, GUY ANTHONY
;  APPLICANT:  PATEL, SARASWATI R.
;  TITLE OF INVENTION:  METHODS OF TREATMENT USING NOVEL LIGANDS
;  TITLE OF INVENTION:  OF THE NEUROPEPTIDE RECEPTOR HFGAN72 AND AGONISTS OR
;  TITLE OF INVENTION:  ANTAGONISTS THEREOF
;  FILE REFERENCE:  P50745
;  CURRENT APPLICATION NUMBER: US/09/211,823C
;  CURRENT FILING DATE:  1998-12-15
;  PRIOR APPLICATION NUMBER:  US 60/069,459
;  PRIOR FILING DATE:  1997-12-15
;  PRIOR APPLICATION NUMBER:  US 60/069,785
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-211-823C-22

  Query Match             99.8%;  Score 2214;  DB 4;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 1.8e-192;
  Matches  423;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
```

```
Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 4
US-08-846-704-4
; Sequence 4, Application US/08846704
; Patent No. 6020157
;  GENERAL INFORMATION:
;    APPLICANT:  BERGSMA, DERK J.
;    APPLICANT:  ELLIS, CATHERINE E.
;    TITLE OF INVENTION:  NOVEL G-PROTEIN COUPLED
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/846,704
;      FILING DATE:  30-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  402 amino acids
```

```
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-846-704-4

  Query Match             95.0%;  Score 2108;  DB 3;  Length 402;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-183;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402


RESULT 5
US-08-462-509B-2
; Sequence 2, Application US/08462509B
; Patent No. 6410701
;  GENERAL INFORMATION:
;    APPLICANT:  Soppet, Daniel et al
;    TITLE OF INVENTION:  Human Neuropeptide Receptor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockiville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/462,509B
;       FILING DATE:  05-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/US95/05616
;       FILING DATE:  05-MAY-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wales, Michele M.
;       REGISTRATION NUMBER:  43,975
;       REFERENCE/DOCKET NUMBER:  PF168P1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-309-8504
;       TELEFAX:  301-309-8439
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;      LENGTH: 402 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-462-509B-2

  Query Match             94.8%;  Score 2104;  DB 4;  Length 402;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-182;
  Matches  401;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
```

```
RESULT 6
PCT-US95-05616-2
; Sequence 2, Application PC/TUS9505616
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Neuropeptide Receptor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05616
;      FILING DATE:  concurrently
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-268
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  402 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
PCT-US95-05616-2

  Query Match             94.8%;  Score 2104;  DB 5;  Length 402;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-182;
  Matches  401;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
                |||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402


RESULT 7
US-08-846-705-2
; Sequence 2, Application US/08846705
; Patent No. 5935814
;  GENERAL INFORMATION:
;    APPLICANT:  BERGSMA, DERK J.
;    APPLICANT:  ELLIS, CATHERINE E
;    TITLE OF INVENTION:  NOVEL G-PROTEIN COUPLED
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/846,705
;      FILING DATE:  30-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-846-705-2

  Query Match             85.7%;  Score 1902.5;  DB 2;  Length 389;
  Best Local Similarity   96.1%;  Pred. No. 2.6e-164;
  Matches  368;  Conservative    2;  Mismatches   10;  Indels    3;  Gaps    2;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSC-CLPGLGP 382
              |||   ::   | ||   ||  |
Db        361 LSG--CKEKSLALSCPSCPGHDP 381


RESULT 8
US-09-211-823C-23
; Sequence 23, Application US/09211823C
; Patent No. 6664229
; GENERAL INFORMATION:
;  APPLICANT:  HAGEN, JAMES JOSEPH
;  APPLICANT:  TERRETT, JONATHAN ALEXANDER
;  APPLICANT:  UPTON, NEIL
;  APPLICANT:  PIPER, DAVID
;  APPLICANT:  SMITH, MARTIN IAN
;  APPLICANT:  KENNETT, GUY ANTHONY
;  APPLICANT:  PATEL, SARASWATI R.
;  TITLE OF INVENTION:  METHODS OF TREATMENT USING NOVEL LIGANDS
;  TITLE OF INVENTION:  OF THE NEUROPEPTIDE RECEPTOR HFGAN72 AND AGONISTS OR
;  TITLE OF INVENTION:  ANTAGONISTS THEREOF
;  FILE REFERENCE:  P50745
;  CURRENT APPLICATION NUMBER: US/09/211,823C
;  CURRENT FILING DATE:  1998-12-15
;  PRIOR APPLICATION NUMBER:  US 60/069,459
;  PRIOR FILING DATE:  1997-12-15
```

```
;  PRIOR APPLICATION NUMBER:  US 60/069,785
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;   LENGTH: 389
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-211-823C-23

  Query Match             85.7%;  Score 1902.5;  DB 4;  Length 389;
  Best Local Similarity   96.1%;  Pred. No. 2.6e-164;
  Matches  368;  Conservative    2;  Mismatches   10;  Indels    3;  Gaps    2;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSC-CLPGLGP 382
              |||   ::   | ||   ||  |
Db        361 LSG--CKEKSLALSCPSCPGHDP 381


RESULT 9
PCT-US95-05616-6
; Sequence 6, Application PC/TUS9505616
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Neuropeptide Receptor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
```

```
;       ZIP:  07068
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 INCH DISKETTE
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WORD PERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/05616
;       FILING DATE:  concurrently
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  FERRARO, GREGORY D.
;       REGISTRATION NUMBER:  36,134
;       REFERENCE/DOCKET NUMBER:  325800-268
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-994-1700
;       TELEFAX:  201-994-1744
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  377 BASE PAIRS
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
;     MOLECULE TYPE:  cDNA
PCT-US95-05616-6

  Query Match             85.7%;  Score 1901;  DB 5;  Length 377;
  Best Local Similarity   96.6%;  Pred. No. 3.4e-164;
  Matches  366;  Conservative    2;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPG 379
              |||   :    : ||  ||
Db        361 LSGCKEKSLVLSPSC--PG 377
```

RESULT 10
US-08-462-509B-4
; Sequence 4, Application US/08462509B
; Patent No. 6410701
; GENERAL INFORMATION:
;   APPLICANT: Soppet, Daniel et al
;   TITLE OF INVENTION: Human Neuropeptide Receptor
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockiville
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/462,509B
;   FILING DATE: 05-JUN-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: WO PCT/US95/05616
;   FILING DATE: 05-MAY-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Wales, Michele M.
;   REGISTRATION NUMBER: 43,975
;   REFERENCE/DOCKET NUMBER: PF168P1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-309-8504
;   TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 369 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-462-509B-4

  Query Match 85.5%; Score 1897; DB 4; Length 369;
  Best Local Similarity 99.7%; Pred. No. 7.6e-164;
  Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
                |||
Db        361 LSG 363
```

```
RESULT 11
PCT-US95-05616-4
; Sequence 4, Application PC/TUS9505616
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Neuropeptide Receptor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05616
;      FILING DATE:  concurrently
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-268
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  369 BASE PAIRS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
```

; MOLECULE TYPE: cDNA
PCT-US95-05616-4

  Query Match 85.5%; Score 1897; DB 5; Length 369;
  Best Local Similarity 99.7%; Pred. No. 7.6e-164;
  Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
              |||
Db        361 LSG 363
```

RESULT 12
US-08-462-509B-6
; Sequence 6, Application US/08462509B
; Patent No. 6410701
; GENERAL INFORMATION:
;   APPLICANT: Soppet, Daniel et al
;   TITLE OF INVENTION: Human Neuropeptide Receptor
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockiville
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:

```
;      APPLICATION NUMBER:  US/08/462,509B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US95/05616
;      FILING DATE:  05-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wales, Michele M.
;      REGISTRATION NUMBER:  43,975
;      REFERENCE/DOCKET NUMBER:  PF168P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  372 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-462-509B-6

  Query Match             85.4%;  Score 1894;  DB 4;  Length 372;
  Best Local Similarity   99.4%;  Pred. No. 1.4e-163;
  Matches  361;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSRDPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
              |||
Db        361 LSG 363


RESULT 13
US-09-426-290-2
; Sequence 2, Application US/09426290
```

; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-426-290-2

Query Match 65.8%; Score 1460; DB 4; Length 444;
Best Local Similarity 68.8%; Pred. No. 3.6e-124;
Matches 284; Conservative 50; Mismatches 69; Indels 10; Gaps 5;

```
Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:|||:||:||||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              ||||||||||||||||||||  ||||:|:|||||:| || :|||||||| |||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| |||||||||||||:|||||:||| ||: || || ||:||| |||||:| | |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVIIWIVSCIIMIPQAIVMECSTVFPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||||  ::|||:|| |||:|||:||| || :|| ||||||| |||||||:| :
Db        204 TTLFTVCDERWGGEIYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263

Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|  |     :|     : |  |  ||:||:|||||||:|||||||||||:||||
Db        264 RKWKPLQPVSQ----PRGPGQPTKSRMSAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||:||||||||||    ||| ||| ||||||||||||||||||||||||:|||||
Db        320 ISILNVLKRVFGMFAHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAF 379

Qy        374 SCCLPGLGPCGSLKAPSPRSSA-SHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              |||  |:      : |:|  | |||: |  :  ::||:|| |||||::|:
Db        380 SCCCLGVHHRQEDRLTRGRTSTESRKSLTTQISNFDNISKLSEQVVLTSISTL 432
```

RESULT 14
US-09-119-788-2
; Sequence 2, Application US/09119788
; Patent No. 6166193
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; TITLE OF INVENTION: CDNA CLONE MY1 THAT ENCODES
; TITLE OF INVENTION: A NOVEL HUMAN 7-TRANSMEMBRANE RECEPTOR

```
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  United States of America
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/119,788
;      FILING DATE:  21-JUL-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/053,790
;      FILING DATE:  25-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GH50029
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5515
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  444 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-119-788-2

  Query Match             65.8%;  Score 1459;  DB 3;  Length 444;
  Best Local Similarity   68.8%;  Pred. No. 4.4e-124;
  Matches  284;  Conservative   50;  Mismatches   69;  Indels   10;  Gaps    5;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:|||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              |||||||||||||||||||||  ||||:|:|||||:| || :|||||||| |||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| |||||||||||||:|||||:||| ||: || ||  ||:||| |||||:| | |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVIIWIVSCIIMIPQAIVMECSTVFPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||||  ::|||:|| |||:|||:||| || :|| ||||||| ||||||:| :
Db        204 TTLFTVCDERWGGEIYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263
```

```
Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|  |     :|     : |  |  ||:||:|||||||:||||||||||:||||
Db        264 RKWKPLQPVSQ----PRGPGQPTKSRMGAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||:||||||||||    ||| ||| |||||||||||||||||||||||||||:|||||
Db        320 ISILNVLKRVFGMFAHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAF 379

Qy        374 SCCLPGLGPCGSLKAPSPRSSA-SHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              |||  |:      :    |:|  | |||: |  :  ::||:|| |||||::|:
Db        380 SCCCLGVHHRQEDRLTRGRTSTESRKSLTTQISNFDNISKLSEQVVLTSISTL 432
```

```
RESULT 15
US-08-513-974B-54
; Sequence 54, Application US/08513974B
; Patent No. 6114139
;  GENERAL INFORMATION:
;    APPLICANT:  Hinuma, Shuji
;    APPLICANT:  Hosoya, Masaki
;    APPLICANT:  Fujii, Ryo
;    APPLICANT:  Ohtaki, Tetsuya
;    APPLICANT:  Fukusumi, Shoji
;    APPLICANT:  Ohgi, Kazuhiro
;    TITLE OF INVENTION:  G PROTEIN COUPLED RECEPTOR PROTEIN,
;    TITLE OF INVENTION:  PRODUCTION, AND USE THEREOF
;    NUMBER OF SEQUENCES:  380
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/513,974B
;      FILING DATE:  14-SEP-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP95/01599
;      FILING DATE:  10-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-093989
;      FILING DATE:  19-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-057186
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-007177
;      FILING DATE:  20-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-326611
```

```
;      FILING DATE:  28-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-270017
;      FILING DATE:  02-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-236357
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-236356
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189274
;      FILING DATE:  11-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189273
;      FILING DATE:  11-AUG-1945
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189272
;      FILING DATE:  11-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resnick, David S.
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45753
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  263 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-513-974B-54

  Query Match             59.1%;  Score 1311;  DB 3;  Length 263;
  Best Local Similarity   96.6%;  Pred. No. 5.8e-111;
  Matches  254;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         91 ADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICH 150
              |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db          1 ADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVVVLTLSSIALDRWYAICH 60

Qy        151 PLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERWADDL 210
              |||||||||||||||||||||||:||||||||||||||||||| ||||||||||
Db         61 PLLFKSTARRARGSILGIWAVSLAVMVPQAAVMECSSVLPELANRTRLLSVCDERWADDL 120

Qy        211 YPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQ 270
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |  |
Db        121 YPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLDDQGQ 180

Qy        271 GLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQA 330
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GLSSEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQA 240

Qy        331 SDREAVYACFTFSHWLVYANSAA 353
              |||||:||||||||||||||||
```

Db 241 SDREAIYACFTFSHWLVYANSAA 263

Search completed: September 28, 2004, 10:01:35
Job time : 36 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 28, 2004, 09:40:21 ; Search time 40 Seconds
(without alignments)
1022.035 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	464.5	20.9	427	2	S50150	gastric CCK-A rece
2	447.5	20.2	449	2	A41738	neuropeptide Y rec
3	434	19.6	381	2	I39187	neuropeptide Y/pep
4	432.5	19.5	370	1	I52315	G protein-coupled
5	428.5	19.3	428	2	JN0692	cholecystokinin ty
6	421.5	19.0	444	2	A42685	cholecystokinin re
7	419	18.9	407	2	S23510	neurokinin 1 recep
8	418	18.8	407	2	A34357	neurokinin 1 recep
9	417.5	18.8	519	2	S17783	tachykinin recepto
10	417	18.8	407	1	JQ1274	neurokinin 1 recep
11	417	18.8	407	2	S20304	neurokinin 1 recep
12	417	18.8	430	2	I51898	cholecystokinin A
13	416	18.7	465	1	JQ1517	neurokinin 3 recep

14	414.5	18.7	452	2	A34916	neurokinin 3 recep
15	413.5	18.6	402	2	I56595	neurokinin 2 recep
16	413	18.6	436	2	JC5599	cholecystokinin-A
17	409.5	18.5	452	2	JC2459	gastrin/cholecysto
18	408.5	18.4	385	2	S55524	neurokinin 3 recep
19	406.5	18.3	450	2	JQ1614	gastrin receptor -
20	403.5	18.2	398	1	JQ1059	neurokinin 2 recep
21	403.5	18.2	452	2	A46195	cholecystokinin B
22	403	18.2	384	1	S00516	neurokinin 2 recep
23	401	18.1	584	2	JC7809	sulfakinin recepto
24	400.5	18.0	423	2	B40470	glucocorticoid-ind
25	400.5	18.0	440	2	A44081	kappa-type opioid
26	399	18.0	447	2	A47430	gastrin/cholecysto
27	398	17.9	349	2	I59336	galanin receptor 1
28	397	17.9	384	2	I57957	neurokinin 2 recep
29	396	17.8	390	2	A36737	neurokinin 2 recep
30	394	17.8	453	2	S32817	gastrin receptor -
31	393.5	17.7	423	2	JC7677	allatostatin recep
32	389	17.5	504	2	A41783	tachykinin recepto
33	385.5	17.4	399	2	S29480	bombesin receptor
34	384	17.3	366	2	S71152	neuropeptide Y/pep
35	383.5	17.3	384	2	S20303	neurokinin 2 recep
36	382.5	17.2	443	2	D40470	glucocorticoid-ind
37	381.5	17.2	384	2	A39003	bombesin/gastrin-r
38	379	17.1	477	1	QRHUB1	beta-1-adrenergic
39	378.5	17.1	390	2	B41007	bombesin receptor,
40	378	17.0	394	2	JC7209	galanin receptor -
41	372.5	16.8	387	2	JC5949	galanin receptor 2
42	372.5	16.8	480	2	I53053	beta 1 adrenergic
43	368.5	16.6	375	2	S63685	neuropeptide Y rec
44	368.5	16.6	384	2	I57682	bombesin/ GRP rece
45	368.5	16.6	399	2	A46632	bombesin-like pept

ALIGNMENTS

RESULT 1
S50150
gastric CCK-A receptor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Apr-2000
C;Accession: S50150
R;Reuben, M.; Rising, L.; Prinz, C.; Hersey, S.; Sachs, G.
Biochim. Biophys. Acta 1219, 321-327, 1994
A;Title: Cloning and expression of the rabbit gastric CCK-A receptor.
A;Reference number: S50150; MUID:95002144; PMID:7918628
A;Accession: S50150
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <REU>
C;Superfamily: neurokinin 1 receptor

Query Match 20.9%; Score 464.5; DB 2; Length 427;
Best Local Similarity 29.6%; Pred. No. 8e-32;
Matches 128; Conservative 85; Mismatches 168; Indels 51; Gaps 13;

```
Qy          8 GAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEW---VLIAAYVAVFVVALVGN 64
              |   |:||      |     ::| |  | :    |   ||   | |  |  :|::::::||
Db          9 GNASGIPP------PCELGLDNETLFCLDQP---PPSKEWQPAVQILLYSLIFLLSVLGN 59

Qy         65 TLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPY 124
              |||   : ||  |||||| |:::|:::|:::   |:| :|: :: : ::|| ||||   |
Db         60 TLVITVLIRNKRMRTVTNIFLLSLAISDLMLCLFCMPFNLIPNLLKDFIFGSALCKTTTY 119

Qy        125 LQAVSVSVAVLTLSFIALDRWYAICHPLLFK--STARRARGSILGIWAVSLAIMVPQAAV 182
              |   ||||: | |  |:|:|: ||| ||  :   |   |   |   | :| ||| |
Db        120 LMGTSVSVSTLNLVAISLERYGAICKPLQSRVWQTKSHALKVIAATWCLSFAIMTPYPIY 179

Qy        183 MECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKL 242
                 |:::|      :  ::|      |:  : :|:   :: :| |  :| :||  |  :|
Db        180 ---SNLVPFTKTNNQTANMCRFLLPSDVMQQAWHTFLLLILFLIPGIVMMVAYGMISLEL 236

Qy        243 W-GRQIPGTTSALVRNWK-------------------RPSDQLGDLEQGLSGEPQPRARA 282
              : | :   :     :  |                   :|: || :|:| |||    |
Db        237 YQGIKFDASQKKSAKERKASTGSGRFEDNDGCYLQRSKPTRQL-ELQQ-LSGGGGGRVSR 294

Qy        283 F--LAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACF 340
                  :    : |:::  :||||:::::| ||::||   |  :    :   :::|
Db        295 IHSSSSAAALMAKKRVIRMLMVIVVLFFLCWMPIFSANAWRAYDTV---SAERRLSGTPI 351

Qy        341 TFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSL 400
              :|   | | :|  ||||| |:: :||  | | | || |  ||      | ||: |  :
Db        352 SFILLLSYTSSCVNPIIYCFMNKRFRLGFMATFPCC-PNPGP------PGPRAEAGEEEE 404

Qy        401 SLQSRCSVSKIS 412
                :| |:|: |
Db        405 GRTTRASLSRYS 416
```

RESULT 2
A41738
neuropeptide Y receptor - fruit fly (Drosophila melanogaster)
N;Alternate names: G protein-coupled receptor PR4
C;Species: Drosophila melanogaster
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 20-Apr-2000
C;Accession: A41738
R;Li, X.J.; Wu, Y.N.; North, R.A.; Forte, M.
J. Biol. Chem. 267, 9-12, 1992
A;Title: Cloning, functional expression, and developmental regulation of a neuropeptide Y receptor from Drosophila melanogaster.
A;Reference number: A41738; MUID:92112730; PMID:1370455
A;Accession: A41738
A;Molecule type: mRNA
A;Residues: 1-449 <LIA>
A;Cross-references: GB:M81490; NID:g157996; PIDN:AAA28727.1; PID:g157997
C;Genetics:
A;Gene: FlyBase:NepYr
A;Cross-references: FlyBase:FBgn0004842
C;Superfamily: neurokinin 1 receptor
C;Keywords: appetite; G protein-coupled receptor; transmembrane protein

Query Match 20.2%; Score 447.5; DB 2; Length 449;
Best Local Similarity 29.4%; Pred. No. 2.4e-30;

```
  Matches  109;  Conservative   63;  Mismatches  138;  Indels   61;  Gaps    9;

Qy         26 DYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFI 85
              ||:      |  |      :: ::   |: :|: ||:||  ||  |:    |||||||||
Db         75 DYD-----LLSEDMWSSAYFKIIVYMLYIPIFIFALIGNGTVCYIVYSTPRMRTVTNYFI 129

Qy         86 VNLSLADVLVTAICLPASLL-VDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDR 144
               :|:: |:|::  | |:| : : |   | || |||  : | ||||| |:  ||  |::||
Db        130 ASLAIGDILMSFFCEPSSFISLFILNYWPFGLALCHFVNYSQAVSVLVSAYTLVAISIDR 189

Qy        145 WYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDE 204
               : ||  ||  : | | |   | |:| ::||  :|   |      :     :   : :| |
Db        190 YIAIMWPLKPRITKRYATFIIAGVWFIALATALPIPIVSGLDIPMSPWHTKCEKY-ICRE 248

Qy        205 RWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQ 264
               |        |    | : :: |||::   | :|  ::| :: ||         :   ||
Db        249 MWPSRSQEYYYTLSLFALQFVVPLGVLIFTYARITIRVWAKRPPGEA-------ETNRDQ 301

Qy        265 LGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVF 324
                                  :  |::||  ||:: |::||  |:|| ::| :|
Db        302 ------------------------RMARSKRKMVKMMLTVVIVFTCCWLPFNILQLL---- 334

Qy        325 GMFRQASDREAV------YACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFS---- 374
                  :| |        |  | | |||  ::   ||||| ::: :||  |
Db        335 -----LNDEEFAHWDPLPYVWFAF-HWLAMSHCCYNPIIYCYMNARFRSGFVQLMHRMPG 388

Qy        375 ----CCLPGLG 381
                   |||  :|
Db        389 LRRWCCLRSVG 399
```

RESULT 3
I39187
neuropeptide Y/peptide YY receptor Y2 - human
N;Alternate names: neuropeptide y/peptide YY receptor type 2
C;Species: Homo sapiens (man)
C;Date: 01-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 20-Apr-2000
C;Accession: I39187; I39163; G02301
R;Gerald, C.; Walker, M.W.; Vaysse, P.J.
J. Biol. Chem. 270, 26758-26761, 1995
A;Title: Expression cloning and pharmacological characterization of a human hippocampal neuropeptide Y/peptide YY Y2 receptor subtype.
A;Reference number: I39187; MUID:96070760; PMID:7592910
A;Accession: I39187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-381 <GER>
A;Cross-references: EMBL:U36269; NID:g1063633; PIDN:AAC50281.1; PID:g1063634
R;Rose, P.M.; Fernandes, P.; Lynch, J.S.; Frazier, S.T.; Fisher, S.M.; Kodukula, K.; Kienzle, B.; Seethala, R.
J. Biol. Chem. 270, 22661-22664, 1995
A;Title: Cloning and functional expression of a cDNA encoding a human type 2 neuropeptide Y receptor.
A;Reference number: I39163; MUID:96032678; PMID:7559383
A;Accession: I39163
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA

A;Residues: 1-133,'A',135-381 <ROS>
A;Cross-references: EMBL:U32500; NID:g1000750; PIDN:AAA93170.1; PID:g1000751
R;Yan, H.; Yang, J.; Marasco, J.; Yamaguchi, K.; Brenner, S.; Collins, F.; Karbon, W.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01019
A;Accession: G02301
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-171,'G',173,'R',175-201,'P',203-208,'A',210-381 <YAN>
A;Cross-references: EMBL:U42389; NID:g1314329; PIDN:AAB07760.1; PID:g1314330
C;Genetics:
A;Gene: GDB:NPY2R
A;Cross-references: GDB:4365607; OMIM:162642
A;Map position: 4q31-4q31
C;Superfamily: neurokinin 1 receptor
C;Keywords: appetite; G protein-coupled receptor; glycoprotein; lipoprotein; thiolester bond; transmembrane protein
F;49-76/Domain: transmembrane #status predicted <TM1>
F;87-113/Domain: transmembrane #status predicted <TM2>
F;166-186/Domain: transmembrane #status predicted <TM4>
F;221-237/Domain: transmembrane #status predicted <TM5>
F;269-291/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;123-203/Disulfide bonds: #status predicted
F;342/Binding site: palmitate (Cys) (covalent) #status predicted
F;372/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.6%; Score 434; DB 2; Length 381;
Best Local Similarity 27.6%; Pred. No. 2.8e-29;
Matches 112; Conservative 74; Mismatches 140; Indels 80; Gaps 10;

```
Qy          3 PSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALV 62
              |  ||  ::             || | | :     |     : : ||| || :: :: ::
Db         24 PQTTPRGEL------------VPDPEPELI-----DSTKLIEVQVVLILAYCSIILLGVI 66

Qy         63 GNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVI 122
              ||:||   | :   ||||||:|| ||::||:||  :||| :|   :   |  |  || ::
Db         67 GNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLTYTLMGEWKMGPVLCHLV 126

Qy        123 PYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGI-WAVSLAIMVPQAA 181
              || | ::| |: :||: |||||   | : |  |  ::|    |:|: | :|  :  | |
Db        127 PYAQGLAVQVSTITLTVIALDRHRCIVYHLESK-ISKRISFLIIGLAWGISALLASPLAI 185

Qy        182 VMECS--SVLPELANRTRLFSVCDERWADD---LYPKIYHSCFFIVTYLAPLGLMAMAYF 236
                | |   ::|:          | |:|  :   :|  :|     :: |: |||::: :|
Db        186 FREYSLIEIIPDFE-----IVACTEKWPGEEKSIYGTVYSLSSLLILYVLPLGIISFSYT 240

Qy        237 QIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKT 296
              :|: ||     ||                                      |     : |:||
Db        241 RIWSKLKNHVSPGA----------------------------------ANDHYHQRRQKT 266

Qy        297 AKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPI 356
               |||: |::|||: :||   |:  :    :  |  | :     ||  | :   :: |||:
Db        267 TKMLVCVVVVFAVSWLP---LHAFQLAVDIDSQVLDLKEYKLIFTVFHIIAMCSTFANPL 323

Qy        357 IYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
```

```
              :| :::  :|: | :|| |                |  | |  :|:
Db        324 LYGWMNSNYRKAFLSAFRC--------------EQRLDAIHSEVSV 355


RESULT 4
I52315
G protein-coupled receptor UHR-1 - rat
C;Species: Rattus sp. (rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-May-2000
C;Accession: I52315
R;Welch, S.K.; O'Hara, B.F.; Kilduff, T.S.; Heller, H.C.
Biochem. Biophys. Res. Commun. 209, 606-613, 1995
A;Title: Sequence and tissue distribution of a candidate G-coupled receptor
cloned from rat hypothalamus.
A;Reference number: I52315; MUID:95251659; PMID:7733930
A;Accession: I52315
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-370 <RES>
A;Cross-references: GB:S77867; NID:g998527; PIDN:AAB34129.1; PID:g998528
C;Superfamily: neurokinin 1 receptor

  Query Match             19.5%;  Score 432.5;  DB 1;  Length 370;
  Best Local Similarity   32.2%;  Pred. No. 3.7e-29;
  Matches  106;  Conservative   66;  Mismatches  116;  Indels   41;  Gaps    7;

Qy         44 QYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPAS 103
              | : :::  |  | || |||| |: | : |   :  |||: | ||:|:|||: | |:| :
Db         58 QLKGLIVMLYSIVVVVGLVGNCLLVLVIARVRRLHNVTNFLIGNLALSDVLMCAACVPLT 117

Qy        104 LLVDI-TESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRAR 162
              |       |:||  || :: :|| |:| |:| || ||:||:  : |||  : : : :
Db        118 LAYAFEPRGWVFGGGLCHLVFFLQPVTVYVSVFTLTTIAVDRYVVLVHPLRRRISLKLSA 177

Qy        163 GSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERW-ADDLYPKIYHSCFFI 221
               ::|||||:|  : :| |       : |   :  ||   |:| | : :   :||     :
Db        178 YAVLGIWALSAVLALPAAVHTYHVELKP---HDVRL---CEEFWGSQERQRQIYAWGLLL 231

Qy        222 VTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRAR 281
               ||| ||  : ::| ::  ||  | :||: :    :|                   |||
Db        232 GTYLLPLLAILLSYVRVSVKLRNRVVPGSVTQSQADW-----------------DRAR 272

Qy        282 AFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFT 341
                       ||:|  :|:||::|||||:||: : |:|:    :  :| |  |
Db        273 -----------RRRTFCLLVVVVVVFALCWLPLHIFNLLR---DLDPRAIDPYAFGLVQL 318

Qy        342 FSHWLVYANSAANPIIYNFLSGKFREQFK 370
               |||  :::  || || :|   |||: :
Db        319 LCHWLAMSSACYNPFIYAWLHDSFREELR 347


RESULT 5
JN0692
cholecystokinin type A receptor - human
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Apr-2000
C;Accession: JN0692; JN0590
```

R;de Weerth, A.; Pisegna, J.R.; Huppi, K.; Wank, S.A.
Biochem. Biophys. Res. Commun. 194, 811-818, 1993
A;Title: Molecular cloning, functional expression and chromosomal localization of the human cholecystokinin type A receptor.
A;Reference number: JN0692; MUID:93343941; PMID:8343165
A;Accession: JN0692
A;Molecule type: mRNA
A;Residues: 1-428 <DEW>
A;Cross-references: GB:L19315; NID:g306595; PIDN:AAA02819.1; PID:g306596
A;Experimental source: gallbladder
R;Ulrich, C.D.; Ferber, I.; Holicky, E.; Hadac, E.; Buell, G.; Miller, L.J.
Biochem. Biophys. Res. Commun. 193, 204-211, 1993
A;Title: Molecular cloning and functional expression of the human gallbladder cholecystokinin A receptor.
A;Reference number: JN0590; MUID:93277552; PMID:8503909
A;Accession: JN0590
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-428 <ULR>
A;Cross-references: GB:L13605; NID:g306490; PIDN:AAA35659.1; PID:g306491
A;Experimental source: gallbladder
C;Comment: This protein has diverse physiological roles in the gastrointestinal system where it mediates pancreatic growth and enzyme secretion, smooth muscle contraction of the gallbladder and stomach, and secretion from gastric mucosal cells.
C;Genetics:
A;Gene: GDB:CCKAR
A;Cross-references: GDB:141927; OMIM:118444
A;Map position: 4pter-4qter
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein; transmembrane protein
F;40-67/Domain: transmembrane #status predicted <TM1>
F;78-104/Domain: transmembrane #status predicted <TM2>
F;116-137/Domain: transmembrane #status predicted <TM3>
F;158-178/Domain: transmembrane #status predicted <TM4>
F;208-234/Domain: transmembrane #status predicted <TM5>
F;314-332/Domain: transmembrane #status predicted <TM6>
F;350-369/Domain: transmembrane #status predicted <TM7>
F;10,24,190,299/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;245,249,260,290/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;256/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 19.3%; Score 428.5; DB 2; Length 428;
Best Local Similarity 29.2%; Pred. No. 9.4e-29;
Matches 125; Conservative 82; Mismatches 180; Indels 41; Gaps 13;

```
Qy          16 GSREPSPVPPDYEDEFLRYLWRDYLYPKQYEW---VLIAAYVAVFVVALVGNTLVCLAVW 72
               ||    |     |:| |  |  |   | : ||   | |  |  :|:::::|||||   :
Db          11 GSNITPPCELGLENETLFCL--DQPRPSK-EWQPAVQILLYSLIFLLSVLGNTLVITVLI 67

Qy          73 RNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSV 132
               ||  |||||| |:::|:::|:::   |:| :|: :: : ::|| |:||   |    ||||
Db          68 RNKRMRTVTNIFLLSLAVSDLMLCLFCMPFNLIPNLLKDFIFGSAVCKTTTYFMGTSVSV 127
```

```
Qy        133 AVLTLSFIALDRWYAICHPLLFK--STARRARGSILGIWAVSLAIMVPQAAVMECSSVLP 190
              :   |  |:|:|: ||| ||  :   |   |   |   | :|  || |       |:::|
Db        128 STFNLVAISLERYGAICKPLQSRVWQTKSHALKVIAATWCLSFTIMTPYPIY---SNLVP 184

Qy        191 ELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLW------- 243
                |  :  ::|     :|:  : :|:   :: :| |  :| :||  |  :|:
Db        185 FTKNNNQTANMCRFLLPNDVMQQSWHTFLLLILFLIPGIVMMVAYGLISLELYQGIKFEA 244

Qy        244 ------GRQIPGTTSA---------LVRNWKRPSD-QLGDLEQGLSGEPQPRARAFLAEV 287
                     : | |||:          ::  : |   :|  |  | |     | |:  :
Db        245 SQKKSAKERKPSTTSSGKYEDSDGCYLQKTRPPRKLELRQLSTGSSSRAN-RIRS-NSSA 302

Qy        288 KQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLV 347
              : |:::  :||:|::::| ||::||   |  :        :::|       :|   |
Db        303 ANLMAKKRVIRMLIVIVVLFFLCWMPIFSANAWR---AYDTASAERRLSGTPISFILLLS 359

Qy        348 YANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCS 407
              | :|  ||||| |:: :||  | | | || |  || |:                || || |
Db        360 YTSSCVNPIIYCFMNKRFRLGFMATFPCC-PNPGPPGARGEVGEEEEGGTTGASL-SRFS 417

Qy        408 VSKISEHV 415
               | :|  |
Db        418 YSHMSASV 425
```

RESULT 6
A42685
cholecystokinin receptor type A - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Apr-2000
C;Accession: A42685; JC4225; PC2213
R;Wank, S.A.; Harkins, R.; Jensen, R.T.; Shapira, H.; de Weerth, A.; Slattery, T.
Proc. Natl. Acad. Sci. U.S.A. 89, 3125-3129, 1992
A;Title: Purification, molecular cloning, and functional expression of the cholecystokinin receptor from rat pancreas.
A;Reference number: A42685; MUID:92212981; PMID:1313582
A;Accession: A42685
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-444 <WAN>
A;Cross-references: GB:M88096; NID:g203383; PIDN:AAA40899.1; PID:g203384
A;Experimental source: pancreas
A;Note: sequence extracted from NCBI backbone (NCBIN:93814, NCBIP:93815)
R;Takata, Y.; Takiguchi, S.; Funakoshi, A.; Kono, A.
Biochem. Biophys. Res. Commun. 213, 958-966, 1995
A;Title: Gene structure of rat cholecystokinin type-A receptor.
A;Reference number: JC4225; MUID:95382845; PMID:7654260
A;Accession: JC4225
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-223 <TAK>
A;Cross-references: DDBJ:D50608; NID:g1100752
R;Mantamadiotis, T.; Baldwin, G.S.
Biochem. Biophys. Res. Commun. 201, 1382-1389, 1994
A;Title: The seventh transmembrane domain of gastrin/CCK receptors contributes to non-peptide antagonist binding.

A;Reference number: PC2213; MUID:94296413; PMID:8024583
A;Accession: PC2213
A;Status: preliminary
A;Molecule type: protein
A;Residues: 366-389 <MAN>
C;Comment: This G-protein-coupled receptor is present in the gastrointestinal system, vagus nerve and localized areas of the central nervous system. It mediates pancreatic growth and enzyme secretion, smooth muscle contraction of the gallbladder and stomach. It is capable of activating phospholipase C and stimulating trunsduction by increasing levels of diacylglycerol, inositol phosphate, and inducing the subsequent release of intracellular calcium.
C;Genetics:
A;Gene: CCKAR
A;Introns: 53/1; 137/1; 224/2; 267/1
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F;57-82/Domain: transmembrane #status predicted <TM1>
F;93-119/Domain: transmembrane #status predicted <TM2>
F;131-151/Domain: transmembrane #status predicted <TM3>
F;173-193/Domain: transmembrane #status predicted <TM4>
F;225-249/Domain: transmembrane #status predicted <TM5>
F;330-348/Domain: transmembrane #status predicted <TM6>
F;366-389/Domain: transmembrane #status predicted <TM7>

Query Match 19.0%; Score 421.5; DB 2; Length 444;
Best Local Similarity 28.1%; Pred. No. 3.9e-28;
Matches 120; Conservative 83; Mismatches 171; Indels 53; Gaps 12;

```
Qy         16 GSREPSPVPPDYEDEFLRYLWRDYLYPKQYEW---VLIAAYVAVFVVALVGNTLVCLAVW 72
              ||    |     |:| |  |  |   | : ||   : |  |  :|:::::|||||   :
Db         26 GSNITPPCELGLENETLFCL--DQPQPSK-EWQSALQILLYSIIFLLSVLGNTLVITVLI 82

Qy         73 RNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSV 132
              ||  |||||| |:::|:::|:::   |:| :|: :: : ::|| |:||   |    ||||
Db         83 RNKRMRTVTNIFLLSLAVSDLMLCLFCMPFNLIPNLLKDFIFGSAVCKTTTYFMGTSVSV 142

Qy        133 AVLTLSFIALDRWYAICHPLLFK--STARRARGSILGIWAVSLAIMVPQAAVMECSSVLP 190
              :   |  |:|:|: ||| ||  :   |   |   |   | :|  || |       |:::|
Db        143 STFNLVAISLERYGAICRPLQSRVWQTKSHALKVIAATWCLSFTIMTPYPIY---SNLVP 199

Qy        191 ELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLW-GRQIPG 249
                 |  :  ::|      |   : : :   :: :| |  :| :||  |  :|: | :
Db        200 FTKNNNQTANMCRFLLPSDAMQQSWQTFLLLILFLLPGIVMVVAYGLISLELYQGIKFDA 259

Qy        250 TTSALVRNWKRPSD----------------------QLGDLEQGLSGEPQPRARAFLAE 286
              :   :  |:||                       :|  |  |  |    | |:  :
Db        260 SQKKSAKE-KKPSTGSSTRYEDSDGCYLQKSRPPRKLELQQLSSGSGGSRLNRIRS-SSS 317

Qy        287 VKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWL 346
                : |:::  :||:|::::| ||::||   |  :   :  ::::       :|   |
Db        318 AANLIAKKRVIRMLIVIVVLFFLCWMPIFSANAWRAYDTV---SAEKHLSGTPISFILLL 374

Qy        347 VYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCG----------SLKAPSPRSS 394
               | :|  ||||| |:: :||  | | | || |  || |           :::|   | |
Db        375 SYTSSCVNPIIYCFMNKRFRLGFMATFPCC-PNPGPPGVRGEVGEEEDGRTIRALLSRYS 433

Qy        395 ASHKSLS 401
```

```
                 || | |
Db        434 YSHMSTS 440
```

RESULT 7
S23510
neurokinin 1 receptor - guinea pig
N;Alternate names: NK-1 receptor; substance P receptor
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 19-May-2000
C;Accession: S23510; S19198
R;Gorbulev, V.; Akhundova, A.; Luzius, H.; Fahrenholz, F.
Biochim. Biophys. Acta 1131, 99-102, 1992
A;Title: Molecular cloning of substance P receptor cDNA from guinea-pig uterus.
A;Reference number: S23510; MUID:92256498; PMID:1374648
A;Accession: S23510
A;Molecule type: mRNA
A;Residues: 1-407 <GOR>
A;Cross-references: EMBL:X64323; NID:g49565; PIDN:CAA45608.1; PID:g49566
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;32-55/Domain: transmembrane #status predicted <TM1>
F;69-89/Domain: transmembrane #status predicted <TM2>
F;117-128/Domain: transmembrane #status predicted <TM3>
F;149-169/Domain: transmembrane #status predicted <TM4>
F;196-217/Domain: transmembrane #status predicted <TM5>
F;250-280/Domain: transmembrane #status predicted <TM6>
F;289-308/Domain: transmembrane #status predicted <TM7>

  Query Match   18.9%; Score 419; DB 2; Length 407;
  Best Local Similarity   27.6%; Pred. No. 5.7e-28;
  Matches 110; Conservative   71; Mismatches 145; Indels   72; Gaps   11;

```
Qy          42 PKQY---EWVLI---AAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLV 95
               | |:    | ::   |||  : | ::||| :|   :  :  ||||||||:|||: |:  :
Db          22 PNQFVQPAWQIVLWAAAYTVIVVTSVVGNVVVMWIILAHKRMRTVTNYFLVNLAFAEASM 81

Qy          96 TAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFK 155
                |     :    :   | :|   ||   :    :|  :: ::: :| ||: || |||  :
Db          82 AAFNTVVNFTYAVHNEWYYGLFYCKFHNFFPIAAVFASIYSMTAVAFDRYMAIIHPLQPR 141

Qy         156 STARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERW---ADDLYP 212
               :|   :  |  || ::| :  ||       |    :  |     ||   |    | :|
Db         142 LSATATKVVICVIWVLALLLAFPQGYY----STTETMPGRV----VCMIEWPSHPDKIYE 193

Qy         213 KIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGL 272
               |:|| |  :: |  || ::  ||  :   ||  :||| :|
Db         194 KVYHICVTVLIYFLPLLVIGYAYTVVGITLWASEIPGDSSDRYH---------------- 237

Qy         273 SGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASD 332
                          :|: |:||  ||::||:  ||:|:||  :  :|  :
Db         238 ---------------EQVSAKRKVVKMMIVVVCTFAICWLPFHIFFLLPYINPDLYLKKF 282

Qy         333 REAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPR 392
               : ||       ||  :::  |||||  |: :||  || || || | :
Db         283 IQQVYLAIM---WLAMSSTMYNPIIYCCLNDRFRLGFKHAFRCC-PFI------------ 326
```

```
Qy        393 SSASHKSLSL------QSRCSVSKISEHVVLTSVTTVL 424
              |:| :: | :      |:: || |:|   : |:::||:
Db        327 SAADYEGLEMKSTRYFQTQGSVYKVSR--LETTISTVV 362
```

RESULT 8
A34357
neurokinin 1 receptor - rat
N;Alternate names: NK-1 receptor; substance P receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 19-May-2000
C;Accession: A38692; A34357; A40089
R;Hershey, A.D.; Dykema, P.E.; Krause, J.E.
J. Biol. Chem. 266, 4366-4374, 1991
A;Title: Organization, structure, and expression of the gene encoding the rat substance P receptor.
A;Reference number: A38692; MUID:91154239; PMID:1705552
A;Accession: A38692
A;Molecule type: DNA
A;Residues: 1-407 <HER>
A;Cross-references: GB:M34751
R;Yokota, Y.; Sasai, Y.; Tanaka, K.; Fujiwara, T.; Tsuchida, K.; Shigemoto, R.; Kakizuka, A.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 264, 17649-17652, 1989
A;Title: Molecular characterization of a functional cDNA for rat substance P receptor.
A;Reference number: A34357; MUID:90036822; PMID:2478537
A;Accession: A34357
A;Molecule type: mRNA
A;Residues: 1-407 <YOK>
A;Cross-references: GB:J05097; NID:g207051; PIDN:AAA42175.1; PID:g207052
R;Hershey, A.D.; Krause, J.E.
Science 247, 958-962, 1990
A;Title: Molecular characterization of a functional cDNA encoding the rat substance P receptor.
A;Reference number: A40089; MUID:90161991; PMID:2154852
A;Accession: A40089
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-74,'R',76-212,'A',214-407 <HE2>
A;Cross-references: GB:M31477
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 18.8%; Score 418; DB 2; Length 407;
Best Local Similarity 27.6%; Pred. No. 7e-28;
Matches 112; Conservative 69; Mismatches 149; Indels 76; Gaps 8;

```
Qy         29 DEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNL 88
              ::|::  |:  |:          |||  : | ::|||:|   :  :  ||||||||:|||
Db         23 NQFVQPTWQIVLW--------AAAYTVIVVTSVVGNVVVIWIILAHKRMRTVTNYFLVNL 74

Qy         89 SLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAI 148
              : |:  : |     :    :   | :|   ||   :    ::  :: ::: :| ||: ||
Db         75 AFAEACMAAFNTVVNFTYAVHNVWYYGLFYCKFHNFFPIAALFASIYSMTAVAFDRYMAI 134

Qy        149 CHPLLFKSTARRARGSILGIWAVSLAIMVPQA---------AVMECSSVLPELANRTRLF 199
```

```
                |||  : :|   :  |  || ::| :  ||          : : |    ||  |||
Db       135 IHPLQPRLSATATKVVIFVIWVLALLLAFPQGYYSTTETMPSRVVCMIEWPEHPNRT--- 191

Qy       200 SVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWK 259
                    | |' || |  :: |  || ::  ||  :   ||  :||| :|
Db       192 -----------YEKAYHICVTVLIYFLPLLVIGYAYTVVGITLWASEIPGDSSDRYH--- 237

Qy       260 RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNV 319
                                      :|: |:|| ||::||:  ||:|:||  |  :
Db       238 ----------------------------EQVSAKRKVVKMMIVVVCTFAICWLPFHVFFL 269

Qy       320 LKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC-LP 378
                |  :          : ||     | ||  ::: |||||  |: :||  || || ||
Db       270 LPYINPDLYLKKFIQQVYLA---SMWLAMSSTMYNPIIYCCLNDRFRLGFKHAFRCCPFI 326

Qy       379 GLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSVTTVL 424
                  |    |:  | |         ||:: || |:|   : |:::||:
Db       327 SAGDYEGLEMKSTR--------YLQTQSSVYKVSR--LETTISTVV 362
```

RESULT 9
S17783
tachykinin receptor homolog DTKR - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000
C;Accession: S17783
R;Li, X.J.; Wolfgang, W.; Wu, Y.N.; North, R.A.; Forte, M.
EMBO J. 10, 3221-3229, 1991
A;Title: Cloning, heterologous expression and developmental regulation of a Drosophila receptor for tachykinin-like peptides.
A;Reference number: S17783; MUID:92007772; PMID:1717263
A;Accession: S17783
A;Molecule type: mRNA
A;Residues: 1-519 <LIX>
A;Cross-references: EMBL:X62711; NID:g8505; PIDN:CAA44595.1; PID:g8506
A;Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in lacking 481-Gly
C;Genetics:
A;Gene: FlyBase:Takr99D
A;Cross-references: FlyBase:FBgn0004622
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; neurotransmitter receptor; transmembrane protein

Query Match 18.8%; Score 417.5; DB 2; Length 519;
Best Local Similarity 28.1%; Pred. No. 1e-27;
Matches 113; Conservative 62; Mismatches 154; Indels 73; Gaps 10;

```
Qy         3 PSATPGAQMGVPPGSREPS-------PVPPDYED--------------EFLRYLWRDYLY 41
               |  |       || |   |       |   | ||               |:   ||  |
Db        45 PCRTLARSSPYPPVSFNHSQTLSTDQPAVGDVEDAAEDAAASMETGSFAFVVPWWRQVL- 103

Qy        42 PKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLP 101
                  | ::  :  : :||  || :|   |     |||||||||||||:|| :|::: :
Db       104 -----WSIL--FGGMVIVATGGNLIVVWIVMTTKRMRTVTNYFIVNLSIADAMVSSLNVT 156

Qy       102 ASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRA 161
```

```
                :    :   | ||   ||:  ::  :|:  :| ||  |::||: ||  ||  : : |
Db        157 FNYYYMLDSDWPFGEFYCKLSQFIAMLSICASVFTLMAISIDRYVAIIRPLQPRMSKRCN 216

Qy        162 RGSILGIWAVSLAIMVPQAAVMECSSV-LPELANRTRLFSVCDERWAD-----DLYPKIY 215
                 ||  |  |  |   :     | :  |:|||    ||   | |          :|
Db        217 LAIAAVIWLASTLISCPMMIIYRTEEVPVRGLSNRT----VCYPEWPDGPTNHSTMESLY 272

Qy        216 HSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGE 275
               :    |:||  |:  | : | ::  :||| :                          ||
Db        273 NILIIILTYFLPIVSMTVTYSRVGIELWGSK------------------------TIGE 307

Qy        276 PQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREA 335
              ||      :|: :|::|:  ||::||:|:||:|:||     ::   :    :|   :
Db        308 CTPR------QVENVRSKRRVVKMMIVVVLIFAICWLPFHSYFIITSCYPAITEAPFIQE 361

Qy        336 VYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCCL 377
              :|      :||  :||  ||||| ::: :||  ||  |  ||
Db        362 LYLAI---YWLAMSNSMYNPIIYCWMNSRFRYGFKMVFRWCL 400
```

RESULT 10
JQ1274
neurokinin 1 receptor - human
N;Alternate names: NK-1 receptor; substance P receptor
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: A41134; JQ1274; JH0478; S21188
R;Gerard, N.P.; Garraway, L.A.; Eddy Jr., R.L.; Shows, T.B.; Iijima, H.; Paquet, J.L.; Gerard, C.
Biochemistry 30, 10640-10646, 1991
A;Title: Human substance P receptor (NK-1): organization of the gene, chromosome localization, and functional expression of cDNA clones.
A;Reference number: A41134; MUID:92031510; PMID:1657150
A;Accession: A41134
A;Molecule type: DNA
A;Residues: 1-328,'G',329-332,334-407 <GER>
A;Cross-references: GB:M76675; NID:g189231
A;Note: in the authors' translation 333-Gly is shown before residue 329 and, consequently, residues 329-332 are displaced one codon to the right
R;Takeda, Y.; Chou, K.B.; Takeda, J.; Sachais, B.S.; Krause, J.E.
Biochem. Biophys. Res. Commun. 179, 1232-1240, 1991
A;Title: Molecular cloning, structural characterization and functional expression of the human substance P receptor.
A;Reference number: JQ1274; MUID:92028856; PMID:1718267
A;Accession: JQ1274
A;Molecule type: mRNA
A;Residues: 1-407 <TAK1>
A;Cross-references: GB:M74290; NID:g338612; PIDN:AAA60601.1; PID:g338613
R;Hopkins, B.; Powell, S.J.; Danks, P.; Briggs, I.; Graham, A.
Biochem. Biophys. Res. Commun. 180, 1110-1117, 1991
A;Title: Isolation and characterization of the human lung NK-1 receptor cDNA.
A;Reference number: JH0478; MUID:92062052; PMID:1659396
A;Accession: JH0478
A;Molecule type: mRNA
A;Residues: 1-407 <HOP>
A;Cross-references: GB:S62045; NID:g237994; PIDN:AAB20168.1; PID:g237995
A;Experimental source: lung

A;Note: the authors translated the codon CAA for residue 31 as Glu
R;Takahashi, K.; Tanaka, A.; Hara, M.; Nakanishi, S.
Eur. J. Biochem. 204, 1025-1033, 1992
A;Title: The primary structure and gene organization of human substance P and neuromedin K receptors.
A;Reference number: S21188; MUID:92201186; PMID:1312928
A;Accession: S21188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <TAK2>
A;Cross-references: GB:X65177; NID:g36636; PIDN:CAA46292.1; PID:g825721
C;Comment: The endogenous ligand of this receptor is neurokinin 1 (substance P), one of the peptides in the mammalian tachykinin system.
C;Genetics:
A;Gene: GDB:TAC1R
A;Cross-references: GDB:128977; OMIM:162323
A;Map position: 2pter-2qter
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane protein
F;32-58/Domain: transmembrane #status predicted <TM1>
F;69-92/Domain: transmembrane #status predicted <TM2>
F;110-128/Domain: transmembrane #status predicted <TM3>
F;149-168/Domain: transmembrane #status predicted <TM4>
F;195-221/Domain: transmembrane #status predicted <TM5>
F;249-273/Domain: transmembrane #status predicted <TM6>
F;286-308/Domain: transmembrane #status predicted <TM7>
F;14,18/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;105-180/Disulfide bonds: #status predicted

Query Match 18.8%; Score 417; DB 1; Length 407;
Best Local Similarity 28.0%; Pred. No. 8.5e-28;
Matches 110; Conservative 69; Mismatches 152; Indels 62; Gaps 10;

```
Qy         42 PKQY---EWVLI---AAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLV 95
              | |:    | ::   |||  : | ::||| :|   :  :  ||||||||:|||: |:  :
Db         22 PNQFVQPAWQIVLWAAAYTVIVVTSVVGNVVVMWIILAHKRMRTVTNYFLVNLAFAEASM 81

Qy         96 TAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFK 155
               |     :    :   | :|   ||   :    :|  :: ::: :| ||: || |||  :
Db         82 AAFNTVVNFTYAVHNEWYYGLFYCKFHNFFPIAAVFASIYSMTAVAFDRYMAIIHPLQPR 141

Qy        156 STARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERWAD---DLYP 212
              :|   :  |  || ::| :  ||       |    : :|     ||   | :    :|
Db        142 LSATATKVVICVIWVLALLLAFPQGYY----STTETMPSRV----VCMIEWPEHPNKIYE 193

Qy        213 KIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGL 272
              |:|| |  :: |  || ::  ||  :   ||  :||| :|
Db        194 KVYHICVTVLIYFLPLLVIGYAYTVVGITLWASEIPGDSSDRYH---------------- 237

Qy        273 SGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASD 332
                          :|: |:||  ||:::||:  ||:|:||  :  :|  :
Db        238 ---------------EQVSAKRKVVKMMIVVVCTFAICWLPFHIFFLLPYINPDLYLKKF 282

Qy        333 REAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC-LPGLGPCGSLKAPSP 391
               : ||       ||  :::  |||||  |: :||  || || ||     |    |: |
Db        283 IQQVYLAIM---WLAMSSTMYNPIIYCCLNDRFRLGFKHAFRCCPFISAGDYEGLEMKST 339
```

```
Qy        392 RSSASHKSLSLQSRCSVSKISEHVVLTSVTTVL 424
                  |          ||:: || |:|    : |:::||:
Db        340 R--------YLQTQGSVYKVSR--LETTISTVV 362
```

RESULT 11
S20304
neurokinin 1 receptor - mouse
N;Alternate names: NK-1 receptor; substance P receptor
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 02-Jun-2000
C;Accession: S20304; I56216; I73044
R;Sundelin, J.B.; Provvedini, D.M.; Wahlestedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson, P.A.
Eur. J. Biochem. 203, 625-631, 1992
A;Title: Molecular cloning of the murine substance K and substance P receptor genes.
A;Reference number: S20303; MUID:92137253; PMID:1370937
A;Accession: S20304
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-407 <SUN>
A;Cross-references: GB:X62934; NID:g54206; PIDN:CAA44707.1; PID:g54207
R;Cook, G.A.; Elliott, D.; Metwali, A.; Blum, A.M.; Sandor, M.; Lynch, R.; Weinstock, J.V.
J. Immunol. 152, 1830-1835, 1994
A;Title: Molecular evidence that granuloma T lymphocytes in murine schistosomiasis mansoni express an authentic substance P (NK-1) receptor.
A;Reference number: I56216; MUID:94165478; PMID:8120392
A;Accession: I56216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 63-290 <COO1>
A;Cross-references: GB:L27826; NID:g450286; PIDN:AAA17891.1; PID:g480776
A;Experimental source: tissue brain
A;Accession: I73044
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 63-290 <COO2>
A;Cross-references: GB:L27828; NID:g450290; PIDN:AAA17892.1; PID:g480778
A;Experimental source: tissue granuloma
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 13.8%; Score 417; DB 2; Length 407;
Best Local Similarity 27.3%; Pred. No. 8.5e-28;
Matches 111; Conservative 70; Mismatches 149; Indels 76; Gaps 8;

```
Qy         29 DEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNL 88
              ::|::  |:  |:          |||  : | ::||| :|   :  :  ||||||||:|||
Db         23 NQFVQPTWQIVLW--------AAAYTVIVVTSVVGNVVVIWIILAHKRMRTVTNYFLVNL 74

Qy         89 SLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAI 148
              : |:  : |     :    :   | :|   ||   :    ::  :: ::: :| ||: ||
Db         75 AFAEACMAAFNTVVNFTYAVHNVWYYGLFYCKFHNFFPIAALFASIYSMTAVAFDRYMAI 134
```

```
Qy        149 CHPLLFKSTARRARGSILGIWAVSLAIMVPQA---------AVMECSSVLPELANRTRLF 199
Db        135 IHPLQPRLSATATKVVIFVIWVLALLLAFPQGYYSTTETMPSRVVCMIEWPEHPNRT--- 191

Qy        200 SVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWK 259
Db        192 -----------YEKAYHICVTVLIYFLPLLVIGYAYTVVGITLWASEIPGDSSDRYH--- 237

Qy        260 RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNV 319
Db        238 ----------------------------EQVSAKRKVVKMMIVVVCTFAICWLPFHIFFL 269

Qy        320 LKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC-LP 378
Db        270 LPYINPDLYLKKFIQQVYLA---SMWLAMSSTMYNPIIYCCLNDRFRLGFKHAFRCCPFI 326

Qy        379 GLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSVTTVL 424
Db        327 SAGDYEGLEMKSTR--------YLQTQSSVYKVSR--LETTISTVV 362
```

RESULT 12
I51898
cholecystokinin A receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
C;Accession: I51898
R;De Weerth, A.; Pisegna, J.R.; Wank, S.A.
Am. J. Physiol. 265, G1116-G1121, 1993
A;Title: Guinea pig gallbladder and pancreas possess identical CCK-A receptor subtypes: receptor cloning and expression.
A;Reference number: I51898; MUID:94106629; PMID:7916580
A;Accession: I51898
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-430 <RES>
A;Cross-references: GB:S68242; NID:g544723; PIDN:AAB29504.1; PID:g544724
C;Superfamily: neurokinin 1 receptor

Query Match 18.8%; Score 417; DB 2; Length 430;
Best Local Similarity 28.4%; Pred. No. 9e-28;
Matches 122; Conservative 81; Mismatches 169; Indels 58; Gaps 13;

```
Qy         16 GSREPSPVPPDYEDEFLRYLWRDYLYPK-QYEW---VLIAAYVAVFVVALVGNTLVCLAV 71
Db         11 GSNITSACELGFENETLFCLDK----PRPSKEWQPAVQILLYSLIFLLSVLGNTLVITVL 66

Qy         72 WRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVS 131
Db         67 IRNKRMRTVTNIFLLSLAVSDLMLCLFCMPFNLIPSLLKDFIFGSAVCKTTTYFMGTSVS 126

Qy        132 VAVLTLSFIALDRWYAICHPLLFK--STARRARGSILGIWAVSLAIMVPQAAVMECSSVL 189
Db        127 VSTFNLVAISLERYGAICKPLQSRVWQTKSHALKVIAATWCLSFTIMTPYPIY---SNLV 183

Qy        190 PELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLW-GRQIP 248
```

```
Db        184 PFTKNNNQTGNMCRFLLPNDVMQQTWHTFLLLILFLIPGIVMMVAYGLISLELYQGIKFD 243

Qy        249 GTTSALVRNWKRPSDQLGDLEQG----LSGEPQPRARAFLAEVKQ--------------- 289
                  :  |  :   | :|      |     ||      |::|
Db        244 AIQKKSAKERKTSTGSSGPMEDSDGCYLQKSRHPRK----LELRQLSPSSSGSNRINRIR 299

Qy        290 -------MRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTF 342
                  : |:::  :||:|::::| ||::||   |  :    :   :::|       :|
Db        300 SSSSTANLMAKKRVIRMLIVIVVLFFLCWMPIFSANAWRAYDTV---SAERHLSGTPISF 356

Qy        343 SHWLVYANSAANPIIYNFLSGKFREQFKAAFSCC----LPGL-GPCG------SLKAPSP 391
               | | :|  ||||| |:: :||  | | | ||     ||: |  |      :  |
Db        357 ILLLSYTSSCVNPIIYCFMNKRFRLGFMATFPCCPNPGTPGVRGEMGEEEEGRTTGASLS 416

Qy        392 RSSASHKSLS 401
              | | || | |
Db        417 RYSYSHMSTS 426
```

RESULT 13
JQ1517
neurokinin 3 receptor - human
N;Alternate names: neuromedin K receptor; NK-3 receptor
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C;Accession: JQ1517; S20435; S21237
R;Huang, R.R.C.; Cheung, A.H.; Mazina, K.E.; Strader, C.D.; Fong, T.M.
Biochem. Biophys. Res. Commun. 184, 966-972, 1992
A;Title: cDNA sequence and heterologous expression of the human neurokinin-3 receptor.
A;Reference number: JQ1517; MUID:92246993; PMID:1374246
A;Accession: JQ1517
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-465 <HUA>
A;Cross-references: GB:M89473; NID:g189223; PIDN:AAA36366.1; PID:g189224
A;Experimental source: brain
R;Buell, G.; Schulz, M.F.; Arkinstall, S.J.; Maury, K.; Missotten, M.; Adami, N.; Talabot, F.; Kawashima, E.
FEBS Lett. 299, 90-95, 1992
A;Title: Molecular characterisation, expression and localisation of human neurokinin-3 receptor.
A;Reference number: S20435; MUID:92183914; PMID:1312036
A;Accession: S20435
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2,'I',4-62,'R',64-465 <BUE>
A;Cross-references: GB:S86392; NID:g246908; PIDN:AAB21706.1; PID:g246909
R;Takahashi, K.; Tanaka, A.; Hara, M.; Nakanishi, S.
Eur. J. Biochem. 204, 1025-1033, 1992
A;Title: The primary structure and gene organization of human substance P and neuromedin K receptors.
A;Reference number: S21188; MUID:92201186; PMID:1312928
A;Accession: S21237
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-438,'F',440-465 <TAK>

A;Cross-references: GB:X65172; NID:g35022; PIDN:CAA46291.1; PID:g825695
C;Comment: The endogenous ligand of this receptor is neurokinin 3 (neuromedin K), one of the peptides in the mammalian tachykinin system.
C;Genetics:
A;Gene: GDB:TACR3
A;Cross-references: GDB:9599126
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane protein
F;83-111/Domain: transmembrane #status predicted <TM1>
F;122-147/Domain: transmembrane #status predicted <TM2>
F;160-181/Domain: transmembrane #status predicted <TM3>
F;202-221/Domain: transmembrane #status predicted <TM4>
F;247-272/Domain: transmembrane #status predicted <TM5>
F;300-321/Domain: transmembrane #status predicted <TM6>
F;333-355/Domain: transmembrane #status predicted <TM7>
F;23,50,73/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;158-233/Disulfide bonds: #status predicted

Query Match 18.7%; Score 416; DB 1; Length 465;
Best Local Similarity 28.7%; Pred. No. 1.2e-27;
Matches 108; Conservative 70; Mismatches 142; Indels 56; Gaps 10;

```
Qy          4 SATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVG 63
              |::| | :|:|  |  ||    :  ::|::  ||  |      | |  ||  |  ||::|
Db         52 SSSPSA-LGLPVASPAPSQPWANLTNQFVQPSWRIAL------WSL--AYGVVVAVAVLG 102

Qy         64 NTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIP 123
              | :|   :  :  ||||||||:|||: :|  : |     : :  :   | ||   |:
Db        103 NLIVIWIILAHKRMRTVTNYFLVNLAFSDASMAAFNTLVNFIYALHSEWYFGANYCRFQN 162

Qy        124 YLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVM 183
              :    :|  :: ::: ||:||: ||  ||  : :|   :  |  || ::  :  ||
Db        163 FFPITAVFASIYSMTAIAVDRYMAIIDPLKPRLSATATKIVIGSIWILAFLLAFPQCLYS 222

Qy        184 ECSSVLPELANRTRLFSVCDERWADDLYPK---IYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              : : |:|    ||    :|  :| :   ||    ||    |: |  || :| : |  :
Db        223 K-TKVMP---GRT----LCFVQWPEG--PKQHFTYHIIVIILVYCFPLLIMGITYTIVGI 272

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
               ||| :||| |                                    :|::|:||  ||:
Db        273 TLWGGEIPGDTCDKYH------------------------------EQLKAKRKVVKMM 301

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ::|:: ||:|:||  :  :|  ::    :    : ||     | ||  :::  |||||
Db        302 IIVVMTFAICWLPYHIYFILTAIYQQLNRWKYIQQVYLA---SFWLAMSSTMYNPIIYCC 358

Qy        361 LSGKFREQFKAAFSCC 376
              |: :||  || ||  |
Db        359 LNKRFRAGFKRAFRWC 374
```

RESULT 14
A34916
neurokinin 3 receptor - rat
N;Alternate names: neuromedin K receptor; NK-3 receptor
C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 19-May-2000
C;Accession: A34916
R;Shigemoto, R.; Yokota, Y.; Tsuchida, K.; Nakanishi, S.
J. Biol. Chem. 265, 623-628, 1990
A;Title: Cloning and expression of a rat neuromedin K receptor cDNA.
A;Reference number: A34916; MUID:90110113; PMID:2153106
A;Accession: A34916
A;Molecule type: mRNA
A;Residues: 1-452 <SHI>
A;Cross-references: GB:J05189; NID:g205670; PIDN:AAA41688.1; PID:g205671
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 18.7%; Score 414.5; DB 2; Length 452;
Best Local Similarity 28.7%; Pred. No. 1.6e-27;
Matches 106; Conservative 66; Mismatches 142; Indels 55; Gaps 8;

```
Qy          11 MGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLA 70
               :|:|  :: || |  :  ::|::  ||  |      | |  ||  |  ||: || :|
Db          45 LGLPATTQAPSQVRANLTNQFVQPSWRIAL------WSL--AYGLVVAVAVFGNLIVIWI 96

Qy          71 VWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSV 130
               :  : ||||||||:|||: :|  | |     : :  :   | ||   |:   :    :|
Db          97 ILAHKRMRTVTNYFLVNLAFSDASVAAFNTLINFIYGLHSEWYFGANYCRFQNFFPITAV 156

Qy         131 SVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLP 190
                 :: ::: ||:||: ||  ||  : :|   :  |  || ::  :  ||       | :
Db         157 FASIYSMTAIAVDRYMAIIDPLKPRLSATATKIVIGSIWILAFLLAFPQCLY----SKIK 212

Qy         191 ELANRTRLFSVCDERWADDLYPK---IYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQI 247
                :  ||    :|  :| :   ||    ||    |: |  || :| : |  :   ||| :|
Db         213 VMPGRT----LCYVQWPEG--PKQHFTYHIIVIILVYCFPLLIMGVTYTIVGITLWGGEI 266

Qy         248 PGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVF 307
               || |                                   :|::|:||  ||:::|:: |
Db         267 PGDTCDKYH------------------------------EQLKAKRKVVKMMIIVVVTF 295

Qy         308 ALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFRE 367
               |:|:||  |  :|  ::    :    : ||     | ||  :::  |||||  |: :||
Db         296 AICWLPYHVYFILTAIYQQLNRWKYIQQVYLA---SFWLAMSSTMYNPIIYCCLNKRFRA 352

Qy         368 QFKAAFSCC 376
                || ||  |
Db         353 GFKRAFRWC 361
```

RESULT 15
I56595
neurokinin 2 receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
C;Accession: I56595
R;Aharony, D.; Little, J.; Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
J. Recept. Res. 14, 399-421, 1994
A;Title: Isolation and characterization of neurokinin A receptor cDNAs from
guinea-pig lung and rabbit pulmonary artery.
A;Reference number: I56595; MUID:95182423; PMID:7877137

A;Accession: I56595
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-402 <RES>
A;Cross-references: GB:S76253; NID:g913274; PIDN:AAB33553.1; PID:g913275
C;Superfamily: neurokinin 1 receptor

Query Match 18.6%; Score 413.5; DB 2; Length 402;
Best Local Similarity 29.7%; Pred. No. 1.7e-27;
Matches 113; Conservative 64; Mismatches 134; Indels 69; Gaps 12;

```
Qy         47 WVL---IAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPAS 103
              | |     ||:|: :||: ||  |   :  :  |||||||||||:|||: : |     :
Db         31 WQLALWATAYLALVLVAVTGNATVTWIILAHQRMRTVTNYFIVNLALADLCMAAFNAAFN 90

Qy        104 LLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARG 163
               :      | || | |          :: |:: ::: ||:||: || ||   : :|   :
Db         91 FVYASHNIWYFGRAFCYFQNLFPITAMFVSIYSMTAIAIDRYMAIVHPFQPRLSAPSTKA 150

Qy        164 SILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPK---IYHSCFF 220
               | ||| |:||:  ||       | : |    |:    |   | :|   |   :||
Db        151 VIGGIWLVALALAFPQCFY----STITEDEGATK----CVVAWPEDSRDKSLLLYHLVVI 202

Qy        221 IVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRA 280
              :: || || :| :||  |   || | :                              ||
Db        203 VLIYLLPLTVMFVAYSIIGLTLWRRAV----------------------------PRH 232

Qy        281 RAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACF 340
              :|  | :: ::|::|  | :::|:: ||:|:||  :  :|          | :| :| |
Db        233 QAHGANLRHLQAKKKFVKTMVLVVVTFAICWLPYHLYFIL---------GSFQEDIY-CH 282

Qy        341 TFSH-------WLVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRS 393
               |         ||  :::  |||||  |: :||  |: || || | : |         :
Db        283 KFIQQVYLALFWLAMSSTMYNPIIYCCLNRRFRSGFRLAFRCC-PWVTP-----TEEDKL 336

Qy        394 SASH-KSLSLQ-SRCSVSKI 411
              :|  | ||: :||   :|
Db        337 ELTHTPSFSLRVNRCHTKEI 356
```

Search completed: September 28, 2004, 10:00:55
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 10:00:18 ; Search time 52 Seconds
(without alignments)
2628.124 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description

1	2218	100.0	425	9	US-09-961-848-2	Sequence 2, Appli
2	2215	99.8	425	9	US-09-828-538-24	Sequence 24, Appl
3	2214	99.8	425	9	US-09-828-538-19	Sequence 19, Appl
4	2214	99.8	425	10	US-09-211-823C-22	Sequence 22, Appl
5	2214	99.8	425	14	US-10-225-567A-368	Sequence 368, App
6	2214	99.8	425	15	US-10-352-684A-22	Sequence 22, Appl
7	2209	99.5	425	11	US-09-826-509-549	Sequence 549, App
8	2108	95.0	402	13	US-10-077-874-2	Sequence 2, Appli
9	2087	94.1	402	10	US-09-393-696-2	Sequence 2, Appli
10	2087	94.1	427	9	US-09-730-931-2	Sequence 2, Appli
11	1902.5	85.7	389	9	US-09-828-538-20	Sequence 20, Appl
12	1902.5	85.7	389	10	US-09-211-823C-23	Sequence 23, Appl
13	1897	85.5	369	13	US-10-077-874-4	Sequence 4, Appli
14	1894	85.4	372	10	US-09-393-696-6	Sequence 6, Appli
15	1894	85.4	372	13	US-10-077-874-6	Sequence 6, Appli
16	1869	84.2	369	10	US-09-393-696-4	Sequence 4, Appli
17	1474.5	66.4	460	12	US-10-081-810-46	Sequence 46, Appl
18	1460	65.8	444	10	US-09-992-331-19	Sequence 19, Appl
19	1460	65.8	444	12	US-10-081-810-45	Sequence 45, Appl
20	1460	65.8	444	14	US-10-225-567A-370	Sequence 370, App
21	1460	65.8	444	14	US-10-262-313-19	Sequence 19, Appl
22	1460	65.8	444	14	US-10-060-369-11	Sequence 11, Appl
23	1460	65.8	444	14	US-10-178-194-2	Sequence 2, Appli
24	1460	65.8	444	16	US-10-768-878-19	Sequence 19, Appl
25	1459	65.8	444	14	US-10-282-717-2	Sequence 2, Appli
26	1455	65.6	444	11	US-09-826-509-551	Sequence 551, App
27	1311	59.1	263	14	US-10-278-087A-54	Sequence 54, Appl
28	522.5	23.5	430	9	US-09-866-248A-8	Sequence 8, Appli
29	522.5	23.5	430	14	US-10-225-567A-658	Sequence 658, App
30	522.5	23.5	430	16	US-10-719-587-54	Sequence 54, Appl
31	522.5	23.5	441	15	US-10-292-798-890	Sequence 890, App
32	519.5	23.4	428	9	US-09-292-973-4	Sequence 4, Appli
33	507.5	22.9	432	9	US-09-866-248A-2	Sequence 2, Appli
34	507.5	22.9	432	16	US-10-719-587-37	Sequence 37, Appl
35	499.5	22.5	420	9	US-09-866-248A-6	Sequence 6, Appli
36	499.5	22.5	420	14	US-10-060-369-9	Sequence 9, Appli
37	499.5	22.5	522	12	US-10-072-012-360	Sequence 360, App
38	499.5	22.5	522	12	US-10-072-012-361	Sequence 361, App
39	499.5	22.5	522	12	US-10-276-774-2093	Sequence 2093, Ap
40	499.5	22.5	522	12	US-10-081-810-53	Sequence 53, Appl
41	499.5	22.5	522	14	US-10-225-567A-512	Sequence 512, App
42	499.5	22.5	522	15	US-10-255-551-2	Sequence 2, Appli
43	499	22.5	417	12	US-10-072-012-358	Sequence 358, App
44	480	21.6	426	9	US-09-292-973-19	Sequence 19, Appl
45	480	21.6	432	9	US-09-292-973-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-961-848-2
; Sequence 2, Application US/09961848
; Patent No. US20020146719A1
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir

```
;  APPLICANT:  Jeffrey Gulcher
;  TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
;  FILE REFERENCE: 2345.1005-004
;  CURRENT APPLICATION NUMBER: US/09/961,848
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/479,128
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/379,083
;  PRIOR FILING DATE: 1999-08-23
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-961-848-2

  Query Match             100.0%;  Score 2218;  DB 9;  Length 425;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-197;
  Matches  424;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 2
US-09-828-538-24
; Sequence 24, Application US/09828538
```

; Patent No. US20010025031A1
; GENERAL INFORMATION:
; APPLICANT: Ellis, Catherine E.
; APPLICANT: Kwok, Cheni
; APPLICANT: Bodsworth, Nicola J.
; APPLICANT: Halsey, Wendy
; APPLICANT: Van Horn, Stephanie
; TITLE OF INVENTION: HFGAN72 Receptor Genomic DNA and Methods
; TITLE OF INVENTION: of Use Thereof in Diagnostic Applications
; FILE REFERENCE: GH-50038-C1
; CURRENT APPLICATION NUMBER: US/09/828,538
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/088,624
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: 60/093,726
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 09/328,014
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 425
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-828-538-24

Query Match 99.8%; Score 2215; DB 9; Length 425;
Best Local Similarity 99.8%; Pred. No. 2e-197;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
```

```
Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 3
US-09-828-538-19
; Sequence 19, Application US/09828538
; Patent No. US20010025031A1
; GENERAL INFORMATION:
; APPLICANT: Ellis, Catherine E.
; APPLICANT: Kwok, Cheni
; APPLICANT: Bodsworth, Nicola J.
; APPLICANT: Halsey, Wendy
; APPLICANT: Van Horn, Stephanie
; TITLE OF INVENTION: HFGAN72 Receptor Genomic DNA and Methods
; TITLE OF INVENTION: of Use Thereof in Diagnostic Applications
; FILE REFERENCE: GH-50038-C1
; CURRENT APPLICATION NUMBER: US/09/828,538
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/088,624
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: 60/093,726
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 09/328,014
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-828-538-19

  Query Match   99.8%; Score 2214; DB 9; Length 425;
  Best Local Similarity   99.5%; Pred. No. 2.5e-197;
  Matches 423; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
```

```
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 4
US-09-211-823C-22
; Sequence 22, Application US/09211823C
; Publication No. US20030087801A1
; GENERAL INFORMATION:
; APPLICANT: HAGEN, JAMES JOSEPH
; APPLICANT: TERRETT, JONATHAN ALEXANDER
; APPLICANT: UPTON, NEIL
; APPLICANT: PIPER, DAVID
; APPLICANT: SMITH, MARTIN IAN
; APPLICANT: KENNETT, GUY ANTHONY
; APPLICANT: PATEL, SARASWATI R.
; TITLE OF INVENTION: METHODS OF TREATMENT USING NOVEL LIGANDS
; TITLE OF INVENTION: OF THE NEUROPEPTIDE RECEPTOR HFGAN72 AND AGONISTS OR
; TITLE OF INVENTION: ANTAGONISTS THEREOF
; FILE REFERENCE: P50745
; CURRENT APPLICATION NUMBER: US/09/211,823C
; CURRENT FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: US 60/069,459
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 60/069,785
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 425
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-211-823C-22

Query Match 99.8%; Score 2214; DB 10; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.5e-197;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
```

```
Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 5
US-10-225-567A-368
; Sequence 368, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS (GPCRS)
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 368
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-368

Query Match 99.8%; Score 2214; DB 14; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.5e-197;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
```

```
Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 6
US-10-352-684A-22
; Sequence 22, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Healy, Aileen
; APPLICANT: Weich, Nadine S.
; APPLICANT: Kelly, Louise M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303,
13906,
; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847,
1849,
; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
; FILE REFERENCE: MPI02-019P1RNOMNIM
; CURRENT APPLICATION NUMBER: US/10/352,684A
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/354,333
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/360,258
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/364,476
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/375,626
; PRIOR FILING DATE: 2002-04-26

```
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-352-684A-22

  Query Match             99.8%;  Score 2214;  DB 15;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 2.5e-197;
  Matches  423;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425
```

RESULT 7
US-09-826-509-549
; Sequence 549, Application US/09826509
; Publication No. US20030204073A1
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 549
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-549

Query Match 99.5%; Score 2209; DB 11; Length 425;
Best Local Similarity 99.3%; Pred. No. 7.4e-197;
Matches 422; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTKKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
```

```
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 8
US-10-077-874-2
; Sequence 2, Application US/10077874
; Publication No. US20020115155A1
;   GENERAL INFORMATION:
;        APPLICANT: Soppet, Daniel et al
;        TITLE OF INVENTION: Human Neuropeptide Receptor
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/077,874
;             FILING DATE: 20-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/462,509
;             FILING DATE: 05-JUNE-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Wales, Michele M.
;             REGISTRATION NUMBER: 43,975
;             REFERENCE/DOCKET NUMBER: PF168P1D1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-309-8504
;             TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 402 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-077-874-2

  Query Match             95.0%;  Score 2108;  DB 13;  Length 402;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-187;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
```

```
Qy          61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy         121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy         181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy         241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy         301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy         361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
               |||||||||||||||||||||||||||||||||||||||||||
Db         361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
```

```
RESULT 9
US-09-393-696-2
; Sequence 2, Application US/09393696
; Publication No. US20030022277A1
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc. et al.
;  TITLE OF INVENTION: Human Neuropeptide Receptor
;  FILE REFERENCE: PF168P2
;  CURRENT APPLICATION NUMBER: US/09/393,696
;  CURRENT FILING DATE: 1999-09-10
;  EARLIER APPLICATION NUMBER: PCT/US95/05616
;  EARLIER FILING DATE: 1995-05-05
;  EARLIER APPLICATION NUMBER: US08/462,509
;  EARLIER FILING DATE: 1995-06-05
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 402
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-393-696-2

  Query Match             94.1%;  Score 2087;  DB 10;  Length 402;
  Best Local Similarity   99.0%;  Pred. No. 1.6e-185;
  Matches  398;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy           1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
               |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db           1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIPAYVAVFVVA 60

Qy          61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy         121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy         181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         181 AVMECSSVLPELANRTRLFSVCHERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy         241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         241 NLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy         301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy         361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
                |||||||||||||||||||||||||||||||||||||||||||
Db         361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
```

RESULT 10
US-09-730-931-2
; Sequence 2, Application US/09730931
; Patent No. US20020064814A1
; GENERAL INFORMATION:
; APPLICANT: ELLIS, CATHERINE E.
; TITLE OF INVENTION: DOG OREXIN 1 RECEPTOR
; FILE REFERENCE: GH-70669
; CURRENT APPLICATION NUMBER: US/09/730,931
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/169,373
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: CANIS FAMILIARIS
US-09-730-931-2

Query Match 94.1%; Score 2087; DB 9; Length 427;
Best Local Similarity 94.1%; Pred. No. 1.7e-185,
Matches 402; Conservative 5; Mismatches 18; Indels 2; Gaps 1;

```
Qy           1 MEPSATPGAQMGVPPGSRE--PSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFV 58
                |||||||||| | | |  |  || ||||||||||||||||||||||||||||||||:
Db           1 MEPSATPGAQTGTPTGGGELSPSLVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFL 60

Qy          59 VALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHAL 118
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          61 VALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHTL 120

Qy         119 CKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVP 178
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||:|||
Db        121 CKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARSSILGIWAVSLAVMVP 180

Qy        179 QAAVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQI 238
                ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 QAAVMECSSVLPELANRTRLFSVCDEHWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQI 240

Qy        239 FRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAK 298
                |||||||||||||||||||||||| |   ||| || ||||||||||||||||||||
Db        241 FRKLWGRQIPGTTSALVRNWKRPSDQLEDQGPGLSAEPPPRARAFLAEVKQMRARRKTAK 300

Qy        299 MLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIY 358
                ||||:|||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        301 MLMVLLLVFALCYLPISVLNVLKRVFGMFRQSSDREAVYACFTFSHWLVYANSAANPIIY 360

Qy        359 NFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLT 418
                |||||||||||||||||||||||||| |||||||||||||| ||||||| |||||||
Db        361 NFLSGKFREQFKAAFSCCLPGLGPCGSPKAPSPRSSASHKSLSLHSRCSVSKVPEHVVLT 420

Qy        419 SVTTVLP 425
                |||||||
Db        421 SVTTVLP 427
```

RESULT 11
US-09-828-538-20
; Sequence 20, Application US/09828538
; Patent No. US20010025031A1
; GENERAL INFORMATION:
; APPLICANT: Ellis, Catherine E.
; APPLICANT: Kwok, Cheni
; APPLICANT: Bodsworth, Nicola J.
; APPLICANT: Halsey, Wendy
; APPLICANT: Van Horn, Stephanie
; TITLE OF INVENTION: HFGAN72 Receptor Genomic DNA and Methods
; TITLE OF INVENTION: of Use Thereof in Diagnostic Applications
; FILE REFERENCE: GH-50038-C1
; CURRENT APPLICATION NUMBER: US/09/828,538
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/088,624
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: 60/093,726
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 09/328,014
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 389
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-828-538-20

Query Match 85.7%; Score 1902.5; DB 9; Length 389;
Best Local Similarity 96.1%; Pred. No. 2.2e-168;
Matches 368; Conservative 2; Mismatches 10; Indels 3; Gaps 2;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSC-CLPGLGP 382
              |||   ::   | ||   ||  |
Db        361 LSG--CKEKSLALSCPSCPGHDP 381
```

RESULT 12
US-09-211-823C-23
; Sequence 23, Application US/09211823C
; Publication No. US20030087801A1
; GENERAL INFORMATION:
; APPLICANT: HAGEN, JAMES JOSEPH
; APPLICANT: TERRETT, JONATHAN ALEXANDER
; APPLICANT: UPTON, NEIL
; APPLICANT: PIPER, DAVID
; APPLICANT: SMITH, MARTIN IAN
; APPLICANT: KENNETT, GUY ANTHONY
; APPLICANT: PATEL, SARASWATI R.
; TITLE OF INVENTION: METHODS OF TREATMENT USING NOVEL LIGANDS
; TITLE OF INVENTION: OF THE NEUROPEPTIDE RECEPTOR HFGAN72 AND AGONISTS OR
; TITLE OF INVENTION: ANTAGONISTS THEREOF
; FILE REFERENCE: P50745
; CURRENT APPLICATION NUMBER: US/09/211,823C
; CURRENT FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: US 60/069,459
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 60/069,785
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 389
; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-211-823C-23

Query Match 85.7%; Score 1902.5; DB 10; Length 389;
Best Local Similarity 96.1%; Pred. No. 2.2e-168;
Matches 368; Conservative 2; Mismatches 10; Indels 3; Gaps 2;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSC-CLPGLGP 382
              |||   ::   | ||   ||  |
Db        361 LSG--CKEKSLALSCPSCPGHDP 381
```

RESULT 13
US-10-077-874-4
; Sequence 4, Application US/10077874
; Publication No. US20020115155A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel et al
; TITLE OF INVENTION: Human Neuropeptide Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/077,874

```
;              FILING DATE: 20-Feb-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/462,509
;              FILING DATE: 05-JUNE-1995
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Wales, Michele M.
;              REGISTRATION NUMBER: 43,975
;              REFERENCE/DOCKET NUMBER: PF168P1D1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 301-309-8504
;              TELEFAX: 301-309-8439
;    INFORMATION FOR SEQ ID NO: 4:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 369 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-077-874-4

  Query Match             85.5%;  Score 1897;  DB 13;  Length 369;
  Best Local Similarity   99.7%;  Pred. No. 6.8e-168;
  Matches  362;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
              |||
Db        361 LSG 363


RESULT 14
US-09-393-696-6
; Sequence 6, Application US/09393696
```

; Publication No. US20030022277A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Human Neuropeptide Receptor
; FILE REFERENCE: PF168P2
; CURRENT APPLICATION NUMBER: US/09/393,696
; CURRENT FILING DATE: 1999-09-10
; EARLIER APPLICATION NUMBER: PCT/US95/05616
; EARLIER FILING DATE: 1995-05-05
; EARLIER APPLICATION NUMBER: US08/462,509
; EARLIER FILING DATE: 1995-06-05
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-393-696-6

Query Match 85.4%; Score 1894; DB 10; Length 372;
Best Local Similarity 99.4%; Pred. No. 1.3e-167;
Matches 361; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSRDPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
              |||
Db        361 LSG 363
```

RESULT 15
US-10-077-874-6
; Sequence 6, Application US/10077874
; Publication No. US20020115155A1
; GENERAL INFORMATION:

```
;         APPLICANT: Soppet, Daniel et al
;         TITLE OF INVENTION: Human Neuropeptide Receptor
;         NUMBER OF SEQUENCES: 12
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Human Genome Sciences, Inc.
;              STREET: 9410 Key West Avenue
;              CITY: Rockville
;              STATE: MD
;              COUNTRY: USA
;              ZIP: 20850
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/077,874
;              FILING DATE: 20-Feb-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/462,509
;              FILING DATE: 05-JUNE-1995
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Wales, Michele M.
;              REGISTRATION NUMBER: 43,975
;              REFERENCE/DOCKET NUMBER: PF168P1D1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 301-309-8504
;              TELEFAX: 301-309-8439
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 372 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-077-874-6

  Query Match             85.4%;  Score 1894;  DB 13;  Length 372;
  Best Local Similarity   99.4%;  Pred. No. 1.3e-167;
  Matches  361;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSRDPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
```

```
Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSG 363
              |||
Db        361 LSG 363
```

Search completed: September 28, 2004, 10:18:29
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 09:37:36 ; Search time 120 Seconds
(without alignments)
1117.460 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description

1	2218	100.0	425	4	Q9HBV6	Q9hbv6 homo sapien
2	1287	58.0	364	11	Q8BV78	Q8bv78 mus musculu
3	976	44.0	199	11	Q80T45	Q80t45 mus musculu
4	659.5	29.7	166	6	Q8MJ13	Q8mj13 ovis aries
5	618	27.9	127	6	Q8SPR4	Q8spr4 ovis aries
6	533	24.0	109	6	Q8I010	Q8i010 bos taurus
7	512	23.1	417	11	Q8BKR6	Q8bkr6 mus musculu
8	509	22.9	417	11	Q924H0	Q924h0 mus musculu
9	508.5	22.9	405	11	Q924N0	Q924n0 mus musculu
10	480	21.6	432	11	Q924G9	Q924g9 rattus norv
11	473.5	21.3	758	5	Q7YU49	Q7yu49 drosophila
12	452.5	20.4	449	5	Q9VB87	Q9vb87 drosophila
13	446	20.1	375	13	O57463	O57463 brachydanio
14	446	20.1	517	5	Q9VWR3	Q9vwr3 drosophila
15	444	20.0	540	5	Q9VRM0	Q9vrm0 drosophila
16	443.5	20.0	436	13	Q7T1P8	Q7t1p8 gallus gall
17	438.5	19.8	393	13	Q7T078	Q7t078 fugu rubrip
18	435.5	19.6	381	11	Q9ERC0	Q9erc0 rattus norv
19	434.5	19.6	381	11	Q8BWV1	Q8bwv1 mus musculu
20	433.5	19.5	678	5	Q94736	Q94736 stomoxys ca
21	428	19.3	373	13	O73734	O73734 brachydanio
22	418.5	18.9	397	5	Q9NHA4	Q9nha4 boophilus m
23	417	18.8	407	11	Q8BYR7	Q8byr7 mus musculu
24	416	18.7	598	5	Q9VWQ9	Q9vwq9 drosophila
25	415	18.7	519	5	Q9VAD2	Q9vad2 drosophila
26	414	18.7	453	11	Q8BKF6	Q8bkf6 mus musculu
27	413	18.6	436	11	Q8VCC7	Q8vcc7 mus musculu
28	412	18.6	374	13	Q9YHX1	Q9yhx1 gadus morhu
29	412	18.6	411	13	Q9W6I3	Q9w6i3 gallus gall
30	412	18.6	504	5	Q9VGX8	Q9vgx8 drosophila
31	412	18.6	517	5	Q8T0S8	Q8t0s8 drosophila
32	411.5	18.5	450	11	P89005	P89005 praomys nat
33	411	18.5	402	5	Q964E5	Q964e5 dugesia tig
34	411	18.5	436	11	Q9DBV6	Q9dbv6 mus musculu
35	409.5	18.5	365	11	Q8BHH0	Q8bhh0 mus musculu
36	406.5	18.3	440	11	Q925R4	Q925r4 cavia porce
37	406	18.3	431	5	Q8T8D1	Q8t8d1 urechis uni
38	405	18.3	429	5	P92045	P92045 lymnaea sta
39	404.5	18.2	398	4	Q969F8	Q969f8 homo sapien
40	403.5	18.2	422	11	Q8VHD7	Q8vhd7 rattus norv
41	402.5	18.1	398	4	Q96QG0	Q96qg0 homo sapien
42	400.5	18.0	396	11	Q924U1	Q924u1 rattus norv
43	400.5	18.0	398	4	Q8NGQ8	Q8ngq8 homo sapien
44	399	18.0	395	11	Q9Z0T7	Q9z0t7 rattus norv
45	399	18.0	396	4	Q92492	Q92492 homo sapien

ALIGNMENTS

RESULT 1
Q9HBV6
ID Q9HBV6 PRELIMINARY; PRT; 425 AA.
AC Q9HBV6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypocretin receptor-1.
GN HCRTR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20429525; PubMed=10973318;
RA Peyron C., Faraco J., Rogers W., Ripley B., Overeem S., Charnay Y.,
RA Nevsimalova S., Aldrich M., Reynolds D., Albin R., Li R., Hungs M.,
RA Pedrazzoli M., Padigaru M., Kucherlapati M., Fan J., Maki R.,
RA Lammers G.J., Bouras C., Kucherlapati R., Nishino S., Mignot E.;
RT "A mutation in a case of early onset narcolepsy and a generalized
RT absence of hypocretin peptides in human narcoleptic brains.";
RL Nat. Med. 6:991-997(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Yeager M., Welch R., Haque K., Bergen A.;
RT "Genomic sequence of the hypocretin (orexin) receptor 1 (HCRTR1).";
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21580342; PubMed=11723285;
RA Olafsdottir B.R., Rye D.B., Scammell T.E., Matheson J.K.,
RA Stefansson K., Gulcher J.R.;
RT "Polymorphisms in hypocretin/orexin pathway genes and narcolepsy.";
RL Neurology 57:1896-1899(2001).
DR EMBL; AF202084; AAG28020.1; -.
DR EMBL; AF202078; AAG28020.1; JOINED.
DR EMBL; AF202079; AAG28020.1; JOINED.
DR EMBL; AF202080; AAG28020.1; JOINED.
DR EMBL; AF202081; AAG28020.1; JOINED.
DR EMBL; AF202082; AAG28020.1; JOINED.
DR EMBL; AF202083; AAG28020.1; JOINED.
DR EMBL; AY070269; AAL50221.1; -.
DR EMBL; AY062030; AAL47214.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016499; F:orexin receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004059; Orexin_receptor1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PRINTS; PR01521; OREXIN1R.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 425 AA; 47535 MW; B650B37F3A2CA096 CRC64;

Query Match 100.0%; Score 2218; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 6.1e-186;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
                 ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
                 |||||
Db        421 TTVLP 425
```

```
RESULT 2
Q8BV78
ID   Q8BV78      PRELIMINARY;      PRT;   364 AA.
AC   Q8BV78;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Orexin receptor type 2.
GN   MOX2R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK079572; BAC37688.1; -.
DR   MGD; MGI:1889024; Mox2r.
DR   GO; GO:0016021; C:integral to membrane; IEA.
```

```
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   364 AA;  42018 MW;  4EB78485DBFABDB4 CRC64;

  Query Match             58.0%;  Score 1287;  DB 11;  Length 364;
  Best Local Similarity   70.1%;  Pred. No. 1.8e-104;
  Matches  244;  Conservative   40;  Mismatches   54;  Indels   10;  Gaps    4;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:||||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              |||||||||||||||||||||  ||||:|:|||||:| || :||||||||| |||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| ||||||||||||:|||||:||| ||: || ||  ||:||| ||||||:|| |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVVIWIVSCIIMIPQAIVMECSSMLPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||| |  ::|||:|| |||:|||:||| || :|| ||||||| ||||||:| :
Db        204 TTLFTVCDEHWGGEVYPKMYHICFFLVTYMAPLFLMILAYLQIFRKLWCRQIPGTSSVVQ 263

Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|  |     :|   : : |  | ||:||:|||||||:||||||||||:|||
Db        264 RKWKQQQPVSQ----PRGSGQQSKARVSAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFL 361
              ||:||||||||||    ||| ||| ||| ||||||||   |   |:|
Db        320 ISILNVLKRVFGMFTHTEDRETVYAWFTFPHWLVYANSCCKP---NYL 364


RESULT 3
Q80T45
ID   Q80T45      PRELIMINARY;      PRT;   199 AA.
AC   Q80T45;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Orexin receptor type 1 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22584407; PubMed=12679517;
RA   Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA   Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA   Bergmann J.E., Gaitanaris G.A.;
RT   "The G Protein-Coupled Receptor Repertoires of Human and Mouse.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
DR   EMBL; AY255599; AAO85111.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     199    199
SQ   SEQUENCE   199 AA;  22773 MW;  8190589414A81185 CRC64;

  Query Match             44.0%;  Score 976;  DB 11;  Length 199;
  Best Local Similarity   94.5%;  Pred. No. 1.6e-77;
  Matches  188;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        141 ALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFS 200
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 ALDRWYAICHPLLFKSTARRARGSILGIWAVSLAVMVPQAAVMECSSVLPELANRTRLFS 60

Qy        201 VCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKR 260
              |||| |||:|||||||||||||||||||| ||||||||||||||||||||||||||||
Db         61 VCDEHWADELYPKIYHSCFFIVTYLAPLGLMGMAYFQIFRKLWGRQIPGTTSALVRNWKR 120

Qy        261 PSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVL 320
              ||:||    |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        121 PSEQLEAQHQGLCTEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVL 180

Qy        321 KRVFGMFRQASDREAVYAC 339
              ||||||||||||||||||
Db        181 KRVFGMFRQASDREAVYAC 199


RESULT 4
Q8MJ13
ID   Q8MJ13      PRELIMINARY;      PRT;   166 AA.
AC   Q8MJ13;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Orexin receptor 2 (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang S., Blackberry M.A., Blache D.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF532967; AAM97918.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
```

```
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     166    166
SQ   SEQUENCE   166 AA;  19127 MW;  DEAB9D51D14727E1 CRC64;

  Query Match             29.7%;  Score 659.5;  DB 6;  Length 166;
  Best Local Similarity   78.3%;  Pred. No. 6.9e-50;
  Matches  123;  Conservative   17;  Mismatches   16;  Indels    1;  Gaps    1;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:||||||| |: ||||||:|| |||||||:||
Db          9 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCLAVWKNH 68

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              ||||||||||||||||||||  ||||:|:|||||:| || :||||||||| |||||:||
Db         69 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 128

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVS 172
              ||| ||||||||||||:|||||:||| ||: || ||
Db        129 TLSCIALDRWYAICHPLMFKSTAKRARNSIVIIWIVS 165


RESULT 5
Q8SPR4
ID   Q8SPR4      PRELIMINARY;      PRT;   127 AA.
AC   Q8SPR4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Orexin receptor 1 (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang S., Blackberry M.A., Blache D.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF499612; AAM18967.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     127    127
SQ   SEQUENCE   127 AA;  14421 MW;  DC5446C6BFE34417 CRC64;

  Query Match             27.9%;  Score 618;  DB 6;  Length 127;
```

```
  Best Local Similarity   92.9%;  Pred. No. 2.2e-46;
  Matches  118;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        167 GIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLA 226
              ||||||||:||||||||||||||||||| |||||||||||||||||||||||:|||||
Db          1 GIWAVSLAVMVPQAAVMECSSVLPELANRTWLFSVCDERWADDLYPKIYHSCFFVVTYLA 60

Qy        227 PLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAE 286
              |||||||||||||||||||||||||||||||| || |  |||  ||||||||||||
Db         61 PLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSVQLEDQGQGLGAEPQPRARAFLAE 120

Qy        287 VKQMRAR 293
              |||||||
Db        121 VKQMRAR 127


RESULT 6
Q8I010
ID   Q8I010      PRELIMINARY;      PRT;   109 AA.
AC   Q8I010;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Orexin receptor 1 (Fragment).
GN   OX1R.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RA   Watanabe K., Aso H., Sato T., Tahara K., Takano S., Yamaguchi T.;
RT   "Expression of orexin receptor 1 in bovine hypothalamus.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB092488; BAC16765.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     109    109
SQ   SEQUENCE   109 AA;  12442 MW;  E8C239847783926B CRC64;

  Query Match             24.0%;  Score 533;  DB 6;  Length 109;
  Best Local Similarity   92.7%;  Pred. No. 5.2e-39;
  Matches  101;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        169 WAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPL 228
              ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 WAVSLAVMVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPL 60

Qy        229 GLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQ 277
```

```
                ||||||||||||||||||||||||||||||||| || :  |||  |||
Db          61 GLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSVQLEEQGQGLCTEPQ 109


RESULT 7
Q8BKR6
ID   Q8BKR6      PRELIMINARY;      PRT;   417 AA.
AC   Q8BKR6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neuropeptide NPFF receptor.
GN   GPR74.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Body;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK050939; BAC34468.1; -.
DR   MGD; MGI:1860130; Gpr74.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0008188; F:neuropeptide receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005397; NPFF_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01572; NPFFRECEPTR2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   417 AA;  47449 MW;  BB8D85EF405D5786 CRC64;

  Query Match           23.1%;  Score 512;  DB 11;  Length 417;
  Best Local Similarity  32.9%;  Pred. No. 1.6e-36;
  Matches  121;  Conservative   64;  Mismatches  121;  Indels   62;  Gaps    7;

Qy         27 YEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIV 86
              | |  : |:   ||:  |   | |::|: :||: :||||:||  | || || ||||:||:
Db         27 YSDINITYV-NYYLHQPQVAAVFISSYLLIFVLCMVGNTVVCFIVIRNRHMHTVTNFFIL 85

Qy         87 NLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWY 146
              ||:::|:||   |:| :|| :|   | || ::||:   :| :||: :| ||  ||:|||:
Db         86 NLAISDLLVGIFCMPITLLDNIIAGWPFGSSMCKISGLVQGISVAASVFTLVAIAVDRFR 145

Qy        147 AICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFS------ 200
               : :|   | | : |  :|: || :::||| | | ::    |  |   | || |
```

```
Db        146 CVVYPFKPKLTVKTAFVTIVIIWGLAIAIMTPSAIMLH---VQEEKYYRVRLSSHNKTST 202

Qy        201 --VCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNW 258
                 | | |      :|| :  |   ||||| |: : | :|   |:       |:|
Db        203 VYWCREDWPRHEMRRIYTTVLFATIYLAPLSLIVIMYARIGASLF------KTAAHCTGK 256

Qy        259 KRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLN 318
              :||                          :    : ::|  |||: | |:| | :||:  |
Db        257 QRP----------------------VQWHVSKKKQKVIKMLLTVALLFILSWLPLWTLM 293

Qy        319 VLK----------RVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQ 368
              :|           |:  ::            : |:||| : ||: ||||| | :  ||
Db        294 MLSDYTDLSPNKLRIINIY-----------IYPFAHWLAFCNSSVNPIIYGFFNENFRNG 342

Qy        369 FKAAFSCC 376
              |: ||  |
Db        343 FQDAFQIC 350
```

```
RESULT 8
Q924H0
ID   Q924H0      PRELIMINARY;      PRT;   417 AA.
AC   Q924H0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Neuropeptide NPFF receptor.
GN   GPR74.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA   Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA   Wang R., Evans J., Gould R., Austin C.P.;
RT   "Identification and characterization of two cognate receptors for
RT   mammalian FMRFamide-like neuropeptides.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF330054; AAK94198.1; -.
DR   MGD; MGI:1860130; Gpr74.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0008188; F:neuropeptide receptor activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005397; NPFF_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01572; NPFFRECEPTR2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
```

```
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   417 AA;  47485 MW;  AC0686514CDAF40C CRC64;

  Query Match             22.9%;  Score 509;  DB 11;  Length 417;
  Best Local Similarity   32.9%;  Pred. No. 3e-36;
  Matches  121;  Conservative   63;  Mismatches  122;  Indels   62;  Gaps    7;

Qy          27 YEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIV 86
               | |  : |:   ||:  |   | |::|: :||: :||||:||  | || || ||||:||:
Db          27 YSDINITYV-NYYLHQPQVAAVFISSYLLIFVLCMVGNTVVCFIVIRNRHMHTVTNFFIL 85

Qy          87 NLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWY 146
               ||: :|:||   |:| :|| :|   | || ::||:   :| :||: :| ||  ||:||:
Db          86 NLATSDLLVGIFCMPITLLDNIIAGWPFGSSMCKISGLVQGISVAASVFTLVAIAVDRFR 145

Qy         147 AICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFS------ 200
                : :|   | | : |  :|: || :::||| | | ::    |  |   | || |
Db         146 CVVYPFKPKLTVKTAFVTIVIIWGLAIAIMTPSAIMLH---VQEEKYYRVRLSSHNKTST 202

Qy         201 --VCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNW 258
                  | | |      :|| :  |   ||||| |: : | :|   |:       |:|
Db         203 VYWCREDWPRHEMRRIYTTVLFATIYLAPLSLIVIMYARIGASLF------KTAAHCTGK 256

Qy         259 KRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLN 318
               :||                          :    : ::|  |||: | |:| | :||:  |
Db         257 QRPE----------------------QWHVSKKKQKVIKMLLTVALLFILSWLPLWTLM 293

Qy         319 VLK----------RVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQ 368
               :|           |:  ::           : |:||| : ||: ||||| | :  ||
Db         294 MLSDYTDLSPNKLRIINIY-----------IYPFAHWLAFCNSSVNPIIYGFFNENFRNG 342

Qy         369 FKAAFSCC 376
               |: ||  |
Db         343 FQDAFQIC 350


RESULT 9
Q924N0
ID   Q924N0      PRELIMINARY;      PRT;   405 AA.
AC   Q924N0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   G-protein-coupled receptor 74.
GN   GPR74.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker R.M., Herzog H.;
RT   "Molecular cloning and characterisation of GPR74 a novel G-protein
RT   coupled receptor closest related to the Y-receptor family.";
RL   Brain Res. Mol. Brain Res. 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
```

```
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF236084; AAK58514.1; -.
DR   MGD; MGI:1860130; Gpr74.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0008188; F:neuropeptide receptor activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005397; NPFF_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01572; NPFFRECEPTR2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   405 AA;  46248 MW;  8592068E1A17D64B CRC64;

  Query Match             22.9%;  Score 508.5;  DB 11;  Length 405;
  Best Local Similarity   32.6%;  Pred. No. 3.2e-36;
  Matches  120;  Conservative   65;  Mismatches  120;  Indels   63;  Gaps    7;

Qy          27 YEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIV 86
               | |  : |:   ||:  |   | |::|: :||: :|||||:||  | || || ||||: |:
Db          30 YSDINITYV-NYYLHQPQVAAVFISSYLLIFVLCMVGNTVVCFIVIRNRHMHTVTNFLIL 88

Qy          87 NLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWY 146
               ||:::|:||   |:| :|| :|   | || ::||:   :| :||: :| ||  ||:||:
Db          89 NLAISDLLVGIFCMPITLLDNIIAGWPFGSSMCKISGLVQGISVAASVFTLVAIAVDRFR 148

Qy         147 AICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFS------ 200
                : :|   | | : |  :|: || :::||| | | ::    |  |   | || |
Db         149 CVVYPFKPKLTVKTAFVTIVIIWGLAIAIMTPSAIMLH---VQEEKYYRVRLSSHNKTST 205

Qy         201 --VCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNW 258
                  | | |      :|| :  | : ||||| |: : | :|   |:       |:|
Db         206 VYWCREDWPRHEMRRIYTTVLFAIIYLAPLSLIVIMYARIGASLF------KTAAHCTGK 259

Qy         259 KRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLN 318
               :||   :                          : ::|  |||: | |:| | :||:  |
Db         260 QRPVQCM----------------------YQEKQKVIKMLLTVALLFILSWLPLWTLM 295

Qy         319 VLK----------RVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQ 368
               :|          |:  ::            : |:||| : ||: ||||| | :  ||
Db         296 MLSDYTDLSPNKLRIINIY----------IYPFAHWLAFCNSSVNPIIYGFFNENFRNG 344

Qy         369 FKAAFSCC 376
               |: ||  |
Db         345 FQDAFQIC 352


RESULT 10
Q924G9
ID   Q924G9      PRELIMINARY;      PRT;   432 AA.
AC   Q924G9;
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Neuropeptide NPVF receptor.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA   Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA   Wang R., Evans J., Gould R., Austin C.P.;
RT   "Identification and characterization of two cognate receptors for
RT   mammalian FMRFamide-like neuropeptides.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF330056; AAK94200.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0008188; F:neuropeptide receptor activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005396; NPFF_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01571; NPFFRECEPTR1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   432 AA;  48171 MW;  EAFDF052C9B7688A CRC64;

  Query Match             21.6%;  Score 480;  DB 11;  Length 432;
  Best Local Similarity   35.5%;  Pred. No. 1.1e-33;
  Matches  121;  Conservative   57;  Mismatches  121;  Indels   42;  Gaps    8;

Qy         48 VLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVD 107
              : ||||| :|:: :|||||||    | :| ||||||| ||:||:::|: |   |:| :|: :
Db         45 MFIAAYVLIFLLCIVGNTLVYFIVLKNRHMRTVTNMFILNLAVSDLPVGIFCMPTTLVDN 104

Qy        108 ITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILG 167
              :   | | :| ||:   :| :||| :| ||  ||::|:  | ||   | | |:|  :|
Db        105 LITGWPFDNATCKMSGLVQGMSVSASVFTLVAIAVERFRCIVHPFREKLTLRKALFTIAV 164

Qy        168 IWAVSLAIMVPQAAVM----ECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVT 223
              |||::| || | |  :    |    : :  ||:     |   | :    |:| :  |
Db        165 IWALALLIMCPSAVTLTVTREEHHFMLDARNRSYPLYSCWGAWPEKGMRKVYTAVLFAHI 224

Qy        224 YLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAF 283
              || || |: : | :| |||   | ||                  | |:           |
Db        225 YLVPLALIVVMYVRIARKLC--QAPGPAR--------------DTEE-----------AV 257

Qy        284 LAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDRE----AVYAC 339
```

```
                 :  | | :   || :| | | | :||: || :|      : : |: :    :|||
Db        258 AEGGRTSRRRARVVHMLAMVALFFTLSWLPLWVLLLLID----YGELSELQLHLLSVYA- 312

Qy        340 FTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF--SCCLP 378
              |  :||| : :|:|||||| : :  ||  |:|||    | |
Db        313 FPLAHWLAFFHSSANPIIYGYFNENFRRGFQAAFRAQLCWP 353
```

RESULT 11
Q7YU49
ID   Q7YU49   PRELIMINARY;   PRT;   758 AA.
AC   Q7YU49;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RE47636p.
GN   CG10823.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT009988; AAQ22457.1; -.
SQ   SEQUENCE   758 AA; 83441 MW; A2E3BC8D566E009B CRC64;

  Query Match   21.3%; Score 473.5; DB 5; Length 758;
  Best Local Similarity   30.9%; Pred. No. 7.6e-33;
  Matches 133; Conservative   67; Mismatches 141; Indels   89; Gaps   14;

```
Qy         27 YEDEFLRYLWRD----------YLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHH 76
              | |  ||:   |          | :      |   ||: ||:| |:||: |   | |
Db        179 YNDSALRWEQLDGSVDFGFDPLYRHSLAMSMVYCVAYIVVFLVGLIGNSFVIAVVLRAPR 238

Qy         77 MRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLT 136
              ||||||||||||::||:||   ||||:|: :|   |: |  :|| :||:| |||: :| :
Db        239 MRTVTNYFIVNLAIADILVIVFCLPAILIGNIFVPWMLGWLMCKFVPYIQGVSVAASVYS 298

Qy        137 LSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRT 196
              |  ::|||: ||  ||  : | ||||  |:||| ::|   :|            :|
Db        299 LIAVSLDRFIAIWWPLK-QMTKRRARIMIIGIWVIALVTTIPWLLFF-------DLVPAE 350

Qy        197 RLFS----------VCDERWADDLYPKIYHSCFFIVT-YLAPLGLMAMAYFQIFRKLW 243
               :||           :| | |       :|     :|  || |: |: : |  |: |:
Db        351 EVFSDALVSAYSQPQFLCQEVWPPGTDGNLYFLLANLVACYLLPMSLITLCYVLIWIKVS 410

Qy        244 GRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVV 303
               | ||| :           |:  ::|                    ::: |  |||: |
```

```
Db        411 TRSIPGESK---------DAQMDRMQQ--------------------KSKVKVIKMLVAV 441

Qy        304 LLVFALCYLPISVLNVLKRV-FGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLS 362
              :::| | :||: |  :  |: ||      : | :      : ||  :||  |||:|: ::
Db        442 VILFVLSWLPLYV--IFARIKFGSDISQEEFEILKKVMPVAQWLGSSNSCINPILYS-VN 498

Qy        363 GKFREQFKAAF---SCCLPGLGPCGSLK-----------APSPRSSASHKSLSLQSRCS- 407
               |:|  | |     ||       || |:             : :||  |:: | :|  |
Db        499 KKYRRGFAAIIKSRSC-------CGRLRYYDNVAIASSTTSTRKSSHYHQNSSRKSPSSK 551

Qy        408 ---VSKISEH 414
                 || | ||
Db        552 GNAVSYIYEH 561
```

RESULT 12
Q9VB87
ID Q9VB87 PRELIMINARY; PRT; 449 AA.
AC Q9VB87;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG5811 protein.
GN NEPYR OR CG5811.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE003759; AAF56655.2; -.
DR FlyBase; FBgn0004842; NepYr.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.

```
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   449 AA;  51869 MW;  FD8D0B3D70476EC1 CRC64;

  Query Match             20.4%;  Score 452.5;  DB 5;  Length 449;
  Best Local Similarity   29.4%;  Pred. No. 2.9e-31;
  Matches  109;  Conservative   64;  Mismatches  137;  Indels   61;  Gaps    9;

Qy         26 DYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFI 85
              ||:      |  |      :: ::   |: :|: ||:||  ||  |:    |||||||||
Db         75 DYD-----LLSEDMWSSAYFKIIVYMLYIPIFIFALIGNGTVCYIVYSTPRMRTVTNYFI 129

Qy         86 VNLSLADVLVTAICLPASLL-VDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDR 144
               :|:: |:|::  |:|:| : : |   | || |||  : | |||||| |:  ||  |::||
Db        130 ASLAIGDILMSFFCVPSSFISLFILNYWPFGLALCHFVNYSQAVSVLVSAYTLVAISIDR 189

Qy        145 WYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDE 204
              : ||  ||  : | | |   | |:| ::||  :|   |      :     :   : :| |
Db        190 YIAIMWPLKPRITKRYATFIIAGVWFIALATALPIPIVSGLDIPMSPWHTKCEKY-ICRE 248

Qy        205 RWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQ 264
               |        |    | : :: |||::   | :|  ::| :: ||         :   ||
Db        249 MWPSRTQEYYYTLSLFALQFVVPLGVLIFTYARITIRVWAKRPPGEA-------ETNRDQ 301

Qy        265 LGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVF 324
                                  :  |::||  ||:: |::||  |:|| ::| :|
Db        302 ----------------------RMARSKRKMVKMMLTVVIVFTCCWLPFNILQLL---- 334

Qy        325 GMFRQASDREAV------YACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFS---- 374
                  :| |        |  | | |||  ::   ||||| ::: :||  |
Db        335 -----LNDEEFAHWDPLPYVWFAF-HWLAMSHCCYNPIIYCYMNARFRSGFVQLMHRMPG 388

Qy        375 ----CCLPGLG 381
                  |||  :|
Db        389 LRRWCCLRSVG 399


RESULT 13
O57463
ID   O57463      PRELIMINARY;      PRT;   375 AA.
AC   O57463;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Neuropeptide Y /peptide YY receptor Yb.
GN   NPYRYB.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lundell I., Berglund M.M., Starback P., Salaneck S., Gehlert D.R.,
RA   Larhammar D.;
RT   "Cloning and characterization of a novel neuropeptide Y (NPY) receptor
```

```
RT   subtype in the zebrafish.";
RL   DNA Cell Biol. 0:0-0(1997).
DR   EMBL; AF030245; AAB94616.1; -.
DR   ZFIN; ZDB-GENE-980526-208; npyryb.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   375 AA;  42116 MW;  4B0C9551131A14D1 CRC64;

  Query Match             20.1%;  Score 446;  DB 13;  Length 375;
  Best Local Similarity   29.4%;  Pred. No. 8.8e-31;
  Matches  114;  Conservative   70;  Mismatches  136;  Indels   68;  Gaps   10;

Qy          49 LIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDI 108
               || ||  :  | ||||| : : : |   || ||| |||||| :|:||  :||| :::  :
Db          28 LIVAYSTMLAVGLVGNTCLVVVITRQKEMRNVTNIFIVNLSCSDILVCLVCLPVTIIYTL 87

Qy         109 TESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGI 168
                : |: | ||||| |::| :||:|:: ::  |||:|   | ||  :|   | :  ::  |
Db          88 MDRWILGEALCKVTPFVQCMSVTVSIFSMVLIALERHQLIIHPTGWKPVVRHSYLAVAVI 147

Qy         169 WAVSLAIMVPQAAV-MECSSVLPELANRTRLFS---VCDERWADDLYPKIYHSCFFIVTY 224
               | ::  | :|  :  :  :|    |:     ||   :| |:|  :     | :   :  |
Db         148 WIIACFISLPFLSFNILTNSPFHNLSLPFNPFSDHFICIEQWPSEGNRLTYTTTLLLCQY 207

Qy         225 LAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFL 284
                 || |: : ||:|| :|  |:                      |: :   |  |
Db         208 CLPLALILVCYFRIFLRLSRRK--------------------DMVERARGGRQ------- 240

Qy         285 AEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYAC----- 339
                  |: :  ::   ||  ::  ||||:||   |||   :|       : ||:  |
Db         241 ---KKAKGSKRVNAMLASIVAAFALCWLP---LNVFNTIFDW-----NHEAIPVCQHDAI 289

Qy         340 FTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSAS--- 396
               |:  |    |::  ||:|| ||:  |::: |:  | |    ||  | ::  | |: |
Db         290 FSACHLTAMASTCVNPVIYGFLNNNFQKELKSLLSRC-RCWGPAESYES-FPLSTVSTGI 347

Qy         397 ----------------HKSLSLQSRCSV 408
                             ||  ||: : |:
Db         348 TKGSILSNGSASTYQPHKKNSLEQKESI 375


RESULT 14
Q9VWR3
ID   Q9VWR3      PRELIMINARY;      PRT;   517 AA.
AC   Q9VWR3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG6857 protein.
```

GN CCKLR-17D1 OR CG6857.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

```
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AE003509; AAF48875.2; -.
DR   FlyBase; FBgn0030950; CCKLR-17D1.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   517 AA;  57670 MW;  7E7BA42776E69D11 CRC64;

  Query Match           20.1%;  Score 446;  DB 5;  Length 517;
  Best Local Similarity  30.3%;  Pred. No. 1.3e-30;
  Matches  132;  Conservative   76;  Mismatches  148;  Indels   80;  Gaps   18;

Qy          1 MEPSATPGAQMGVP------PGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYV 54
              |: : || :    |      ||:|    |  |        :|            ::  |
Db        142 MDLATTPSSSTLAPAVTVRTPGNRSVVRVSADVP------IW----------VVPCYS 183

Qy         55 AVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLF 114
              |: : |:||| || | : :|  |||:|| |::||:::|:|:   |:| :|:  :   ::|
Db        184 AILLCAVVGNLLVVLTLVQNRRMRTITNVFLLNLAISDILLGVFCMPVTLVGTLLRHFIF 243

Qy        115 GHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKS--TARRARGSILGIWAVS 172
              |  |||:| : || ||:|:  ||  |: :|:|||||||  ::  |   |   |  ||  |
```

```
Db        244 GELLCKLIQFAQAASVAVSSWTLVAISCERYYAICHPLRSRTWQTINHANKIIAIIWLGS 303

Qy        173 LAIMVPQAAVMECSSVLPELANRTRLFSVCDERW-ADDL-YPKIYHSCFFIVTYLAPLGL 230
              |  | | ||    | ::|   :|  |   | |:| || | | : |:   |:   |  | |
Db        304 LVCMTPIAAF---SQLMP--TSRPGL-RKCREQWPADSLNYERAYN--LFLDLALLVLPL 355

Qy        231 MAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQM 290
              :|:::  :|       |  |    :|| :        :  | ||              :
Db        356 LALSFTYLF-------ITRTLYVSMRNERA-------MNFGSSG--------------L 386

Qy        291 RARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAV-YACFTFSHWLVYA 349
               ::::  ||| |::| | :|: |: |:| :  : |     :  | | |   :|   | |:
Db        387 ESKKRVVKMLFVLVLEFFICWTPLYVINTMTMLLG----PTVYEYVGYTSISFLQLLAYS 442

Qy        350 NSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPS---PRSSASHKSLSLQSR- 405
              :|  ||| | |::  ||  |   |     |:  |  | ||     | | :  :||:
Db        443 SSCCNPITYCFMNASFRRAFVDTFK----GMRVCERLCAPCCFWRRRSKNETNLSVAGNS 498

Qy        406 --CSVSKISEHVVLTS 419
                : | :| | :| |
Db        499 IALANSVMSSHTILES 514
```

```
RESULT 15
Q9VRM0
ID   Q9VRM0       PRELIMINARY;      PRT;   540 AA.
AC   Q9VRM0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG10626 protein.
GN   CG10626.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
```

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AE003566; AAF50775.1; -.
DR HSSP; P02699; 1F88.
DR FlyBase; FBgn0035610; CG10626.
DR GO; GO:0016323; C:basolateral plasma membrane; IDA.
DR GO; GO:0042071; F:leucokinin receptor activity; IDA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 540 AA; 60612 MW; 2D22D5E1BDFD47D5 CRC64;

Query Match 20.0%; Score 444; DB 5; Length 540;
Best Local Similarity 27.8%; Pred. No. 2e-30;
Matches 114; Conservative 73; Mismatches 151; Indels 72; Gaps 9;

```
Qy         15 PGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRN 74
              ||: |        | || |     |  | :   :|   |  : :||::|||||   |
Db         14 PGAEE--------EAEFERL----YAAPAEIVALLSIFYGGISIVAVIGNTLVIWVVATT 61

Qy         75 HHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAV 134
               |||||| :| ||: |||::   |:|      : :||     :|   |::||:||:|:|
Db         62 RQMRTVTNMYIANLAFADVIIGLFCIPFQFQAALLQSWNLPWFMCSFCPFVQALSVNVSV 121

Qy        135 LTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLP---- 190
               ||: ||:||  || :||  : |   ::  | ||| ::|   || |       :
Db        122 FTLTAIAIDRHRAIINPLRARPTKFVSKFIIGGIWMLALLFAVPFAIAFRVEELTERFRE 181

Qy        191 --ELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIP 248
               |  | || | : ::  :||      ::  | | || |  ::: | |:  :||| : |
Db        182 NNETYNVTRPFCM-NKNLSDDQLQSFRYTLVF-VQYLVPFCVISFVYIQMAVRLWGTRAP 239
```

```
Qy        249 GTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFA 308
              |                                      ::  :: ::|  |||::|:::|
Db        240 GNAQD-----------------------------SRDITLLKNKKKVIKMLIIVVIIFG 269

Qy        309 LCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQ 368
              ||:||: : |:|   :    : :|   :   :    ||  :||  || ||   : ||: :
Db        270 LCWLPLQLYNIL---YVTIPEINDYHFISIVWFCCDWLAMSNSCYNPFIYGIYNEKFKRE 326

Qy        369 FKAAFSCCLPGLGPCG--------------SLKAPSPRSSASHKSLSLQS 404
              |   |: |      |                 :| | ||: :: |: ::|
Db        327 FNKRFAACF-----CKFKTSMDAHERTFSMHTRASSIRSTYANSSMRIRS 371
```

Search completed: September 28, 2004, 10:00:11
Job time : 124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 09:21:40 ; Search time 13 Seconds
(without alignments)
1702.294 Million cell updates/sec

Title: US-10-070-532-2
Perfect score: 2219
Sequence: 1 MEPSATPGAQMGVPPGSREP.....CSVSKISEHVVLTSVTTVLP 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2214	99.8	425	1	OX1R_HUMAN	O43613 homo sapien
2	2018.5	91.0	416	1	OX1R_RAT	P56718 rattus norv
3	1478.5	66.6	460	1	OX2R_MOUSE	P58308 mus musculu
4	1474.5	66.4	460	1	OX2R_RAT	P56719 rattus norv
5	1469	66.2	444	1	OX2R_CANFA	Q9tup7 canis famil
6	1460	65.8	444	1	OX2R_HUMAN	O43614 homo sapien
7	1261.5	56.8	260	1	OX1R_MOUSE	P58307 mus musculu
8	522.5	23.5	430	1	NFF1_HUMAN	Q9gzq6 homo sapien
9	507.5	22.9	432	1	NFF1_RAT	Q9ep86 rattus norv
10	499.5	22.5	522	1	NFF2_HUMAN	Q9y5x5 homo sapien
11	499	22.5	417	1	NFF2_RAT	Q9eqd2 rattus norv
12	466.5	21.0	427	1	CCKR_RABIT	O97772 oryctolagus
13	451	20.3	382	1	NY2R_PIG	O02836 sus scrofa
14	447.5	20.2	449	1	NYR_DROME	P25931 drosophila
15	437	19.7	86	1	OX1R_PIG	O97661 sus scrofa
16	435	19.6	384	1	NY2R_BOVIN	P79113 bos taurus
17	434.5	19.6	385	1	NY2R_MOUSE	P97295 mus musculu

18	434	19.6	370	1	GP10_HUMAN	P49683 homo sapien
19	434	19.6	381	1	NY2R_HUMAN	P49146 homo sapien
20	432.5	19.5	370	1	GP10_RAT	Q64121 rattus norv
21	431.5	19.4	381	1	NY2R_CAVPO	Q9z2d5 cavia porce
22	431	19.4	385	1	NY2R_CHICK	Q9ddn6 gallus gall
23	428.5	19.3	428	1	CCKR_HUMAN	P32238 homo sapien
24	427.5	19.3	453	1	CCKR_XENLA	P70031 xenopus lae
25	426	19.2	381	1	NY2R_MACMU	Q9gk74 macaca mula
26	421.5	19.0	444	1	CCKR_RAT	P30551 rattus norv
27	420.5	18.9	467	1	NK3R_RABIT	O97512 oryctolagus
28	419	18.9	407	1	NK1R_CAVPO	P30547 cavia porce
29	418.5	18.9	452	1	NK3R_MOUSE	P47937 mus musculu
30	418	18.8	407	1	NK1R_RAT	P14600 rattus norv
31	417.5	18.8	519	1	TLR2_DROME	P30975 drosophila
32	417	18.8	407	1	NK1R_HUMAN	P25103 homo sapien
33	417	18.8	407	1	NK1R_MOUSE	P30548 mus musculu
34	417	18.8	430	1	CCKR_CAVPO	Q63931 cavia porce
35	417	18.8	453	1	GASR_MOUSE	P56481 mus musculu
36	416.5	18.8	422	1	GP83_CANFA	Q9ttq9 canis famil
37	416	18.7	465	1	NK3R_HUMAN	P29371 homo sapien
38	414.5	18.7	452	1	NK3R_RAT	P16177 rattus norv
39	413.5	18.6	402	1	NK2R_CAVPO	Q64077 cavia porce
40	413	18.6	436	1	CCKR_MOUSE	O08786 mus musculu
41	409.5	18.5	452	1	GASR_RABIT	P46627 oryctolagus
42	408.5	18.4	423	1	GP83_HUMAN	Q9nym4 homo sapien
43	406.5	18.3	450	1	GASR_PRANA	P30796 praomys nat
44	404	18.2	346	1	GALR_RAT	Q62805 rattus norv
45	403.5	18.2	398	1	NK2R_HUMAN	P21452 homo sapien

ALIGNMENTS

```
RESULT 1
OX1R_HUMAN
ID   OX1R_HUMAN      STANDARD;      PRT;    425 AA.
AC   O43613;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 1 (Ox1r) (Hypocretin receptor type 1).
GN   HCRTR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98150861; PubMed=9491897;
RA   Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA   Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA   Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA   McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA   Yanagisawa M.;
RT   "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT   and G protein-coupled receptors that regulate feeding behavior.";
RL   Cell 92:573-585(1998).
```

```
RN   [2]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Moderately selective excitatory receptor for orexin-A
CC       and, with a lower affinity, for orexin-B neuropeptide. Seems to be
CC       exclusively coupled to the G(q) subclass of heteromeric G
CC       proteins, which activates the phospholipase C mediated signaling
CC       cascade (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF041243; AAC39601.1; -.
DR   Genew; HGNC:4848; HCRTR1.
DR   MIM; 602392; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0007631; P:feeding behavior; TAS.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
DR   GO; GO:0007268; P:synaptic transmission; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004059; Orexin_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01521; OREXIN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1     46       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     47     67       1 (POTENTIAL).
FT   DOMAIN       68     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    102       2 (POTENTIAL).
FT   DOMAIN      103    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    120    142       3 (POTENTIAL).
FT   DOMAIN      143    164       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    165    185       4 (POTENTIAL).
FT   DOMAIN      186    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    239       5 (POTENTIAL).
FT   DOMAIN      240    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    299    321       6 (POTENTIAL).
```

```
FT   DOMAIN       322    336       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     337    360       7 (POTENTIAL).
FT   DOMAIN       361    425       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   425 AA;  47521 MW;  1634083DE10CA092 CRC64;

  Query Match             99.8%;  Score 2214;  DB 1;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 5.4e-144;
  Matches  423;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRGRAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSISKISEHVVLTSV 420

Qy        421 TTVLP 425
              |||||
Db        421 TTVLP 425


RESULT 2
OX1R_RAT
ID   OX1R_RAT       STANDARD;      PRT;   416 AA.
AC   P56718;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 1 (Ox1r) (Hypocretin receptor type 1).
GN   HCRTR1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98150861; PubMed=9491897;
RA   Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA   Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA   Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA   McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA   Yanagisawa M.;
RT   "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT   and G protein-coupled receptors that regulate feeding behavior.";
RL   Cell 92:573-585(1998).
RN   [2]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Moderately selective excitatory receptor for orexin-A
CC       and, with a lower affinity, for orexin-B neuropeptide. Seems to be
CC       exclusively coupled to the G(q) subclass of heteromeric G
CC       proteins, which activates the phospholipase C mediated signaling
CC       cascade.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Highly expressed in the brain in the
CC       prefrontal cortex, hippocampus, paraventricular thalamus,
CC       ventromedial hypothalamus, arcuate nucleus, dorsal raphe nucleus,
CC       and locus coeruleus. Not detected in the spleen, lung, liver,
CC       skeletal muscle, kidney and testis. Orexin receptor mRNA
CC       expression has also been reported in the adrenal gland, enteric
CC       nervous system, and pancreas.
CC   -!- INDUCTION: By nutritional state, up-regulated by fasting.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF041244; AAC40041.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004059; Orexin_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01521; OREXIN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
```

```
FT   DOMAIN        1     46       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     47     67       1 (POTENTIAL).
FT   DOMAIN       68     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    102       2 (POTENTIAL).
FT   DOMAIN      103    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    120    142       3 (POTENTIAL).
FT   DOMAIN      143    164       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    165    185       4 (POTENTIAL).
FT   DOMAIN      186    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    239       5 (POTENTIAL).
FT   DOMAIN      240    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    299    321       6 (POTENTIAL).
FT   DOMAIN      322    336       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    337    360       7 (POTENTIAL).
FT   DOMAIN      361    416       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   416 AA;  46799 MW;  774DE7A22EA05D18 CRC64;

  Query Match             91.0%;  Score 2018.5;  DB 1;  Length 416;
  Best Local Similarity   92.0%;  Pred. No. 1.1e-130;
  Matches  390;  Conservative    8;  Mismatches   17;  Indels    9;  Gaps    1;

Qy          1 MEPSATPGAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVA 60
              |||||||||| |||  | ||  :|||||||||||||||||||||||||||||||::|
Db          1 MEPSATPGAQPGVPTSSGEPFHLPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFLIA 60

Qy         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCK 120

Qy        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 VIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAVMVPQA 180

Qy        181 AVMECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFR 240
              |||||||||||||||||||||||||:|||||||||| |||||||||| ||||||||
Db        181 AVMECSSVLPELANRTRLFSVCDERWADELYPKIYHSCFFFVTYLAPLGLMGMAYFQIFR 240

Qy        241 KLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKML 300
              |||| ||||||||||||||||:||    |||  ||||||||||||||||||||||||
Db        241 KLWGPQIPGTTSALVRNWKRPSEQLEAQHQGLCTEPQPRARAFLAEVKQMRARRKTAKML 300

Qy        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNF 360

Qy        361 LSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSRCSVSKISEHVVLTSV 420
              ||||||||||||||||||||         | ||| ||||||||||||:|||||||:|
Db        361 LSGKFREQFKAAFSCCLPGLG---------PSSSARHKSLSLQSRCSVSKVSEHVVLTTV 411

Qy        421 TTVL 424
              ||||
Db        412 TTVL 415


RESULT 3
OX2R_MOUSE
```

```
ID   OX2R_MOUSE     STANDARD;      PRT;   460 AA.
AC   P58308; Q8BG12;
DT   16-OCT-2001 (Rel. 40, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN   HCRTR2 OR MOX2R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum, and Hypothalamus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE OF 100-311 FROM N.A.
RC   STRAIN=C57BL/6;
RA   Szendro P.I., Maevers K., Eichele G.;
RT   "Cloning of mouse orexin receptors.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [4]
```

```
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Nonselective, high-affinity receptor for both orexin-A
CC       and orexin-B neuropeptides.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK038551; BAC30039.1; -.
DR   EMBL; AK048781; BAC33457.1; -.
DR   EMBL; AF394597; AAK71327.1; -.
DR   MGD; MGI:1889024; Mox2r.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004060; Orexin_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01522; OREXIN2R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     54       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     55     75       1 (POTENTIAL).
FT   DOMAIN       76     88       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     89    110       2 (POTENTIAL).
FT   DOMAIN      111    127       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    128    150       3 (POTENTIAL).
FT   DOMAIN      151    172       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    173    193       4 (POTENTIAL).
FT   DOMAIN      194    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    247       5 (POTENTIAL).
FT   DOMAIN      248    304       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    305    327       6 (POTENTIAL).
FT   DOMAIN      328    342       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    343    366       7 (POTENTIAL).
FT   DOMAIN      367    460       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     14     14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    201    201       A -> T (IN REF. 2).
FT   CONFLICT    240    240       I -> V (IN REF. 2).
SQ   SEQUENCE   460 AA;  52461 MW;  D62A67C15BA67DCC CRC64;

  Query Match             66.6%;  Score 1478.5;  DB 1;  Length 460;
  Best Local Similarity   69.7%;  Pred. No. 7.4e-94;
  Matches  287;  Conservative   48;  Mismatches   68;  Indels    9;  Gaps    4;
```

```
Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:||||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              |||||||||||||||||||||  ||||:|:||||:| || :||||||||| |||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| ||||||||||||:|||||:||| ||: || ||  ||:||| ||||||:|| |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVVIWIVSCIIMIPQAIVMECSSMLPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||| |  ::|||:|| |||:|||:||| || :|| ||||||| ||||||:| :
Db        204 TTLFTVCDEHWGGEVYPKMYHICFFLVTYMAPLCLMILAYLQIFRKLWCRQIPGTSSVVQ 263

Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|  |     :|   : : |  |  ||:||:|||||||:||||||||||:||||
Db        264 RKWKQQQPVSQ----PRGSGQQSKARISAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||:||||||||||    ||| ||| |||||||||||||||||||||||||:|||||
Db        320 ISILNVLKRVFGMFTHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAF 379

Qy        374 SCCLPGLGPCGSLKAPSPRSSASHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              ||||      |   |    |: | |||: |  :  :|||:||||||||::|:
Db        380 SCCLGVHHRQGDRLARGRTSTESRKSLTTQISNFDNVSKLSEHVVLTSISTL 431
```

```
RESULT 4
OX2R_RAT
ID   OX2R_RAT       STANDARD;      PRT;   460 AA.
AC   P56719;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN   HCRTR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98150861; PubMed=9491897;
RA   Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA   Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA   Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA   McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA   Yanagisawa M.;
RT   "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT   and G protein-coupled receptors that regulate feeding behavior.";
RL   Cell 92:573-585(1998).
RN   [2]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
```

```
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Nonselective, high-affinity receptor for both orexin-A
CC       and orexin-B neuropeptides.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in the brain in the cerebral
CC       cortex, septal nuclei, hippocampus, medial thalamic groups, dorsal
CC       and median raphe nuclei, and many hypothalamic nuclei including
CC       the tuberomammillary nucleus, dorsomedial hypothalamus,
CC       paraventricular hypothalamic nucleus, and ventral premammillary
CC       nucleus. Not detected in the spleen, lung, liver, skeletal muscle,
CC       kidney and testis. Orexin receptor mRNA expression has also been
CC       reported in the adrenal gland, enteric nervous system, and
CC       pancreas.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041246; AAC40042.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004060; Orexin_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF03827; Orexin_rec2; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01522; OREXIN2R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     54       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     55     75       1 (POTENTIAL).
FT   DOMAIN       76     88       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     89    110       2 (POTENTIAL).
FT   DOMAIN      111    127       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    128    150       3 (POTENTIAL).
FT   DOMAIN      151    172       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    173    193       4 (POTENTIAL).
FT   DOMAIN      194    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    247       5 (POTENTIAL).
FT   DOMAIN      248    304       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    305    327       6 (POTENTIAL).
FT   DOMAIN      328    342       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    343    366       7 (POTENTIAL).
FT   DOMAIN      367    460       CYTOPLASMIC (POTENTIAL).
```

```
FT   CARBOHYD     14      14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22      22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202     202       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   460 AA;  52489 MW;  3B44E3D82F8B85D5 CRC64;

  Query Match             66.4%;  Score 1474.5;  DB 1;  Length 460;
  Best Local Similarity   69.8%;  Pred. No. 1.4e-93;
  Matches  286;  Conservative   48;  Mismatches   71;  Indels    5;  Gaps    3;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:||||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              ||||||||||||||||||||  ||||:|:||||:| || :|||||||| |||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| |||||||||||||:|||||:||| ||: || ||  ||:||| ||| ||:|| |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVVIWIVSCIIMIPQAIVMERSSMLPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||||  ::|||:|| |||:|||:||| || :|| ||||||| ||||||:| :
Db        204 TTLFTVCDERWGGEVYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263

Qy        256 RNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPIS 315
              | ||:|  |     :|   : : |  |  ||:||:|||||||:||||||||||:||||||
Db        264 RKWKQP--QPVSQPRGSGQQSKARISAVAAEIKQIRARRKTARMLMVVLLVFAICYLPIS 321

Qy        316 VLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSC 375
              :||||||||||    ||| ||| |||||||||||||||||||||||||:|||||||
Db        322 ILNVLKRVFGMFTHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAFSC 381

Qy        376 CLPGLGPCGSLKAPSPRSSASHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              ||      |   |    |: | |||: |  :  :|||:|||| |||::|:
Db        382 CLGVHRRQGDRLARGRTSTESRKSLTTQISNFDNVSKLSEHVALTSISTL 431


RESULT 5
OX2R_CANFA
ID   OX2R_CANFA     STANDARD;      PRT;   444 AA.
AC   Q9TUP7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN   HCRTR2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99385793; PubMed=10458611;
RA   Lin L., Faraco J., Li R., Kadotani H., Rogers W., Lin X., Qiu X.,
RA   de Jong P.J., Nishino S., Mignot E.;
RT   "The sleep disorder canine narcolepsy is caused by a mutation in the
```

RT hypocretin receptor 2 gene.";
RL Cell 98:365-376(1999).
RN [2]
RP REVIEW.
RX MEDLINE=21237974; PubMed=11340621;
RA Hungs M., Mignot E.;
RT "Hypocretin/orexin, sleep and narcolepsy.";
RL Bioessays 23:397-408(2001).
RN [3]
RP REVIEW.
RX MEDLINE=21178476; PubMed=11283317;
RA Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT "To eat or to sleep? Orexin in the regulation of feeding and
RT wakefulness.";
RL Annu. Rev. Neurosci. 24:429-458(2001).
RN [4]
RP VARIANT NARCOLEPSY LYS-54, AND MUTAGENESIS OF GLU-54.
RX MEDLINE=21180003; PubMed=11282968;
RA Hungs M., Fan J., Lin L., Lin X., Maki R.A., Mignot E.;
RT "Identification and functional analysis of mutations in the hypocretin
RT (orexin) genes of narcoleptic canines.";
RL Genome Res. 11:531-539(2001).
CC -!- FUNCTION: Nonselective, high-affinity receptor for both orexin-A
CC and orexin-B neuropeptides.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DISEASE: Defects in HCRTR2 are a cause of an autosomal recessive
CC form of narcolepsy, observed in labradors, dobermans and
CC dachshunds. Narcolepsy is a neurological sleep disorder affecting
CC animals and humans, characterized by excessive daytime sleepiness,
CC sleep fragmentation, symptoms of abnormal rapid-eye-mouvement
CC (REM) sleep, such as cataplexy, hypnagogic hallucinations, and
CC sleep paralysis. Cataplexy is a sudden loss of muscle tone
CC triggered by emotions, which is the most valuable clinical feature
CC used to diagnose narcolepsy. As in humans, most cases of canine
CC narcolepsy are sporadic but an autosomal recessive form was also
CC observed.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC ----------

CC ----------
DR EMBL; AF164626; AAD49333.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004060; Orexin_receptor2.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF03827; Orexin_rec2; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PRINTS; PR01522; OREXIN2R.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Disease mutation.

```
FT   DOMAIN        1     54       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     55     75       1 (POTENTIAL).
FT   DOMAIN       76     88       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     89    110       2 (POTENTIAL).
FT   DOMAIN      111    127       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    128    150       3 (POTENTIAL).
FT   DOMAIN      151    172       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    173    193       4 (POTENTIAL).
FT   DOMAIN      194    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    247       5 (POTENTIAL).
FT   DOMAIN      248    304       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    305    327       6 (POTENTIAL).
FT   DOMAIN      328    342       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    343    366       7 (POTENTIAL).
FT   DOMAIN      367    444       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     14     14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      54     54       E -> K (IN AUTOSOMAL RECESSIVE
FT                                NARCOLEPSY).
FT   MUTAGEN      54     54       E->K: LOSS OF FUNCTION.
SQ   SEQUENCE   444 AA;  50675 MW;  D848A4536D485D6B CRC64;

  Query Match             66.2%;  Score 1469;  DB 1;  Length 444;
  Best Local Similarity   69.2%;  Pred. No. 3.2e-93;
  Matches  286;  Conservative   51;  Mismatches   66;  Indels   10;  Gaps    5;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:|||:||||||| |: ||||||||| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALVGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              |||||||||||||||||||| ||||:|:|||||:| || :|||||||| ||||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| |||||||||||||:|||||:||| ||: || ||  ||:||| |||||::|| |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVIIWIVSCIIMIPQAIVMECSTMLPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||||  ::|||:|| |||:|||:||| || :|| ||||||| |||||||:| :
Db        204 TTLFTVCDERWGGEIYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263

Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|: |     :|   : : |  | ||:||:||||||||:||||||||||:||||
Db        264 RKWKPLQPASQ----PRGPGQQTKSRISAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||:||||||||||    ||| ||| ||||||||||||||||||||||||||:|||||
Db        320 ISILNVLKRVFGMFTHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAF 379

Qy        374 SCCLPGLGPCGSLKAPSPRSSA-SHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              |||  |:      :    |:|  | |||: |  :  :|||:|| |||||::|:
Db        380 SCCCLGVHHRQEDRLTRGRTSTESRKSLTTQISNFDNVSKLSEQVVLTSISTL 432

RESULT 6
```

```
OX2R_HUMAN
ID   OX2R_HUMAN      STANDARD;      PRT;   444 AA.
AC   O43614;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN   HCRTR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98150861; PubMed=9491897;
RA   Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA   Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA   Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA   McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA   Yanagisawa M.;
RT   "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT   and G protein-coupled receptors that regulate feeding behavior.";
RL   Cell 92:573-585(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21580342; PubMed=11723285;
RA   Olafsdottir B.R., Rye D.B., Scammell T.E., Matheson J.K.,
RA   Stefansson K., Gulcher J.R.;
RT   "Polymorphisms in hypocretin/orexin pathway genes and narcolepsy.";
RL   Neurology 57:1896-1899(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [4]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Nonselective, high-affinity receptor for both orexin-A
CC       and orexin-B neuropeptides.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041245; AAC39602.1; -.
```

```
DR   EMBL; AY062031; AAL47215.1; -.
DR   Genew; HGNC:4849; HCRTR2.
DR   MIM; 602393; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0008188; F:neuropeptide receptor activity; TAS.
DR   GO; GO:0007631; P:feeding behavior; TAS.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
DR   GO; GO:0007268; P:synaptic transmission; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004060; Orexin_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF03827; Orexin_rec2; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01522; OREXIN2R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     54       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     55     75       1 (POTENTIAL).
FT   DOMAIN       76     88       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     89    110       2 (POTENTIAL).
FT   DOMAIN      111    127       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    128    150       3 (POTENTIAL).
FT   DOMAIN      151    172       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    173    193       4 (POTENTIAL).
FT   DOMAIN      194    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    247       5 (POTENTIAL).
FT   DOMAIN      248    304       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    305    327       6 (POTENTIAL).
FT   DOMAIN      328    342       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    343    366       7 (POTENTIAL).
FT   DOMAIN      367    444       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     14     14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   444 AA;  50680 MW;  CA0669F0D4224C65 CRC64;

  Query Match             65.8%;  Score 1460;  DB 1;  Length 444;
  Best Local Similarity   68.8%;  Pred. No. 1.3e-92;
  Matches  284;  Conservative   50;  Mismatches   69;  Indels   10;  Gaps    5;

Qy         17 SREPSPVPPDYED-EFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNH 75
              ::||   | ||:| ||||||||:||:||:||||||| |: ||||||:|| |||:|||:||
Db         24 TQEPFLNPTDYDDEEFLRYLWREYLHPKEYEWVLIAGYIIVFVVALIGNVLVCVAVWKNH 83

Qy         76 HMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVL 135
              ||||||||||||||||||||  ||||:|:|||||:| || :|||||||| ||||||:||
Db         84 HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQSLCKVIPYLQTVSVSVSVL 143

Qy        136 TLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANR 195
              ||| |||||||||||||:|||||:||| ||: || ||  ||:||| |||||:| | |||:
Db        144 TLSCIALDRWYAICHPLMFKSTAKRARNSIVIIWIVSCIIMIPQAIVMECSTVFPGLANK 203

Qy        196 TRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALV 255
              | ||:|||||  ::|||:|| |||:|||:||| || :|| ||||||| ||||||:| :
Db        204 TTLFTVCDERWGGEIYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263
```

```
Qy        256 RNWK--RPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              | ||  :|  |     :|     : |  |  ||:||:||||||||:||||||||||:||||
Db        264 RKWKPLQPVSQ----PRGPGQPTKSRMSAVAAEIKQIRARRKTARMLMVVLLVFAICYLP 319

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||:||||||||||    ||| ||| ||||||||||||||||||||||||||:|||||
Db        320 ISILNVLKRVFGMFAHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAF 379

Qy        374 SCCLPGLGPCGSLKAPSPRSSA-SHKSLSLQ--SRCSVSKISEHVVLTSVTTV 423
              |||  |:      :    |:|  | |||: |  :  ::||:|| |||||::|:
Db        380 SCCCLGVHHRQEDRLTRGRTSTESRKSLTTQISNFDNISKLSEQVVLTSISTL 432
```

```
RESULT 7
OX1R_MOUSE
ID   OX1R_MOUSE     STANDARD;      PRT;   260 AA.
AC   P58307;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 1 (Ox1r) (Hypocretin receptor type 1) (Fragment).
GN   HCRTR1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   Szendro P.I., Maevers K., Eichele G.;
RT   "Cloning of mouse orexin receptors.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Moderately selective excitatory receptor for orexin-A
CC       and, with a lower affinity, for orexin-B neuropeptide. Seems to be
CC       exclusively coupled to the G(q) subclass of heteromeric G
CC       proteins, which activates the phospholipase C mediated signaling
CC       cascade (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF394596; AAK71326.1; -.
DR   MGD; MGI:2385650; Hcrtr1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004059; Orexin_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01521; OREXIN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   NON_TER       1      1
FT   TRANSMEM     <1      9       3 (POTENTIAL).
FT   DOMAIN       10     31       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     32     52       4 (POTENTIAL).
FT   DOMAIN       53     83       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     84    106       5 (POTENTIAL).
FT   DOMAIN      107    165       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    166    188       6 (POTENTIAL).
FT   DOMAIN      189    203       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    204    227       7 (POTENTIAL).
FT   DOMAIN      228   >260       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     260    260
SQ   SEQUENCE   260 AA;  29419 MW;  AC4E67176AFC662C CRC64;

  Query Match             56.8%;  Score 1261.5;  DB 1;  Length 260;
  Best Local Similarity   91.4%;  Pred. No. 2.6e-79;
  Matches  246;  Conservative    3;  Mismatches   11;  Indels    9;  Gaps    1;

Qy        134 VLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELA 193
              ||||| |||||||||||||||||||||||||||||||||:||||||||||||||||||
Db          1 VLTLSVIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAVMVPQAAVMECSSVLPELA 60

Qy        194 NRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSA 253
              |||||||||| |||:|||||||||||||||||| ||||||||||||||||||||
Db         61 NRTRLFSVCDEHWADELYPKIYHSCFFIVTYLAPLGLMGMAYFQIFRKLWGRQIPGTTSA 120

Qy        254 LVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 313
              |||||||||:||    |||  ||||||||||||||||||||||||||||||||||||||
Db        121 LVRNWKRPSEQLEAQHQGLCTEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLP 180

Qy        314 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 240

Qy        374 SCCLPGLGPCGSLKAPSPRSSASHKSLSL 402
              ||||||||         | ||| ||||||
Db        241 SCCLPGLG---------PGSSARHKSLSL 260


RESULT 8
NFF1_HUMAN
ID   NFF1_HUMAN     STANDARD;      PRT;   430 AA.
```

AC Q9GZQ6; Q8NGR0; Q96RN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuropeptide FF receptor 1 (RFamide-related peptide receptor OT7T022).
GN NPFF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spinal cord;
RX MEDLINE=20564301; PubMed=11024015;
RA Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
RA Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhlani P.P.,
RA Raddatz R., Yao W.-J., Ogozalek K.L., Boyle N., Kouranova E.V.,
RA Quan Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
RA Borowsky B.;
RT "Identification and characterization of two G protein-coupled
RT receptors for neuropeptide FF.";
RL J. Biol. Chem. 275:39324-39331(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20482175; PubMed=11025660;
RA Hinuma S., Shintani Y., Fukusumi S., Iijima N., Matsumoto Y.,
RA Hosoya M., Fujii R., Watanabe T., Kikuchi K., Terao Y., Yano T.,
RA Yamamoto T., Kawamata Y., Habata Y., Asada M., Kitada C., Kurokawa T.,
RA Onda H., Nishimura O., Tanaka M., Ibata Y., Fujino M.;
RT "New neuropeptides containing carboxy-terminal RFamide and their
RT receptor in mammals.";
RL Nat. Cell Biol. 2:703-708(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA Wang R., Evans J., Gould R., Austin C.P.;
RT "Identification and characterization of two cognate receptors for
RT mammalian FMRFamide-like neuropeptides.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Receptor for NPAF (A-18-F-amide) and NPFF (F-8-F-amide)
CC neuropeptides, also known as morphine-modulating peptides. Can
CC also be activated by a variety of naturally ocurring or synthetic
CC FMRF-amide like ligands. This receptor mediates its action by
CC association with G proteins that activate a phosphatidylinositol-
CC calcium second messenger system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to erroneous

```
CC       gene model prediction.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF268898; AAG41397.1; -.
DR   EMBL; AB040104; BAB17677.1; -.
DR   EMBL; AF330055; AAK94199.1; -.
DR   EMBL; AB065729; BAC05950.1; ALT_SEQ.
DR   MIM; 607448; -.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; NAS.
DR   GO; GO:0008188; F:neuropeptide receptor activity; NAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005396; NPFF_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01571; NPFFRECEPTR1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    101       2 (POTENTIAL).
FT   DOMAIN      102    117       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    118    138       3 (POTENTIAL).
FT   DOMAIN      139    158       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    159    179       4 (POTENTIAL).
FT   DOMAIN      180    214       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    215    235       5 (POTENTIAL).
FT   DOMAIN      236    271       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    272    292       6 (POTENTIAL).
FT   DOMAIN      293    307       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    308    328       7 (POTENTIAL).
FT   DOMAIN      329    430       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    116    203       BY SIMILARITY.
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     18     18       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    376    376       V -> A (IN REF. 3).
SQ   SEQUENCE   430 AA;  47818 MW;  BBB0146281B2B9A0 CRC64;

  Query Match             23.5%;  Score 522.5;  DB 1;  Length 430;
  Best Local Similarity   33.6%;  Pred. No. 9.2e-29;
  Matches  137;  Conservative   56;  Mismatches  150;  Indels   65;  Gaps    9;

Qy          2 EPSATPGAQMGVPPGSREP-------SPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYV 54
```

```
                |||          || |  |       :   |     |  |    | :      : | ||
Db            4 EPSQ---------PPNSSWPLSQNGTNTEATPATNLTFSSY----YQHTSPVAAMFIVAYA 51

Qy           55 AVFVVALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLF 114
                :|:: :|||||||  | :| || |||| ||:||:::|:||   |:| :|: ::   | |
Db           52 LIFLLCMVGNTLVCFIVLKNRHMHTVTNMFILNLAVSDLLVGIFCMPTTLVDNLITGWPF 111

Qy          115 GHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLA 174
                :| ||:   :| :||| :| ||  ||::|:  | ||   | | |:|  :|  |||::|
Db          112 DNATCKMSGLVQGMSVSASVFTLVAIAVERFRCIVHPFREKLTLRKALVTIAVIWALALL 171

Qy          175 IMVPQAAVM----ECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGL 230
                || | |  :    |    : :  ||:     | | | :    ::| :  |   ||||| |
Db          172 IMCPSAVTLTVTREEHHFMVDARNRSYPLYSCWEAWPEKGMRRVYTTVLFSHIYLAPLAL 231

Qy          231 MAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQM 290
                : : | :| |||   | ||                              | |      |: :
Db          232 IVVMYARIARKLC--QAPG--------------------------PAPGGEE-AADPRAS 262

Qy          291 RARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYAN 350
                | | :   ||::| | | | :||:  | :|     :          ||| | |:||| : |
Db          263 RRRARVVHMLVMVALFFTLSWLPLWALLLLIDYGQLSAPQLHLVTVYA-FPFAHWLAFFN 321

Qy          351 SAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHK 398
                |:|||||| : :  ||  |:|||   |             || | |||
Db          322 SSANPIIYGYFNENFRRGFQAAFRARL----------CPRPSGSHK 357
```

```
RESULT 9
NFF1_RAT
ID   NFF1_RAT       STANDARD;      PRT;   432 AA.
AC   Q9EP86;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neuropeptide FF receptor 1 (RFamide-related peptide receptor OT7T022).
GN   NPFF1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Brain stem;
RX   MEDLINE=20482175; PubMed=11025660;
RA   Hinuma S., Shintani Y., Fukusumi S., Iijima N., Matsumoto Y.,
RA   Hosoya M., Fujii R., Watanabe T., Kikuchi K., Terao Y., Yano T.,
RA   Yamamoto T., Kawamata Y., Habata Y., Asada M., Kitada C., Kurokawa T.,
RA   Onda H., Nishimura O., Tanaka M., Ibata Y., Fujino M.;
RT   "New neuropeptides containing carboxy-terminal RFamide and their
RT   receptor in mammals.";
RL   Nat. Cell Biol. 2:703-708(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;
RX   MEDLINE=20564301; PubMed=11024015;
RA   Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
```

```
RA   Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhlani P.P.,
RA   Raddatz R., Yao W.-J., Ogozalek K.L., Boyle N., Kouranova E.V.,
RA   Quan Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
RA   Borowsky B.;
RT   "Identification and characterization of two G protein-coupled
RT   receptors for neuropeptide FF.";
RL   J. Biol. Chem. 275:39324-39331(2000).
CC   -!- FUNCTION: Receptor for NPAF (A-18-F-amide) and NPFF (F-8-F-amide)
CC       neuropeptides, also known as morphine-modulating peptides. Can
CC       also be activated by a variety of naturally ocurring or synthetic
CC       FMRF-amide like ligands. This receptor mediates its action by
CC       association with G proteins that activate a phosphatidylinositol-
CC       calcium second messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed at high levels in the hypothalamus.
CC       Moderate levels found in the midbrain, thalamus, medulla
CC       oblongata, testis, eye, whole brain, cerebral cortex, striatum,
CC       hippocampus, cerebellum, optic nerve, placenta, spinal cord,
CC       pituitary gland and ovary.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC       HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB040103; BAB17676.1; -.
DR   EMBL; AF268901; AAG41400.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005396; NPFF_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01571; NPFFRECEPTR1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    101       2 (POTENTIAL).
FT   DOMAIN      102    117       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    118    138       3 (POTENTIAL).
FT   DOMAIN      139    158       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    159    179       4 (POTENTIAL).
FT   DOMAIN      180    214       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    215    235       5 (POTENTIAL).
FT   DOMAIN      236    273       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    274    294       6 (POTENTIAL).
FT   DOMAIN      295    309       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    310    330       7 (POTENTIAL).
FT   DOMAIN      331    432       CYTOPLASMIC (POTENTIAL).
```

```
FT   DISULFID    116    203       BY SIMILARITY.
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     18     18       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   432 AA;  48323 MW;  827325849444C408 CRC64;

  Query Match             22.9%;  Score 507.5;  DB 1;  Length 432;
  Best Local Similarity   34.7%;  Pred. No. 9.7e-28;
  Matches  135;  Conservative   60;  Mismatches  143;  Indels   51;  Gaps   11;

Qy          2 EPSATPGAQMGVPPGSREPSPVPPDYEDE--FLRYLWRDYLYPKQYEWVLIAAYVAVFVV 59
              |||  |    |  |  :  | |          |  |    | :      : ||||| :|::
Db          4 EPSQPPN---GSWPLGQNGSDVETSMATSLTFSSY----YQHSSPVAAMFIAAYVLIFLL 56

Qy         60 ALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALC 119
               :|||||||  | :| |||||||  ||:||:::|:||   |:| :|: ::   | | :| |
Db         57 CMVGNTLVCFIVLKNRHMRTVTNMFILNLAVSDLLVGIFCMPTTLVDNLITGWPFDNATC 116

Qy        120 KVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQ 179
              |:   :| :||| :| ||  ||:::|:  | ||   | | |:|  :|  |||::| || |
Db        117 KMSGLVQGMSVSASVFTLVAIAVERFRCIVHPFREKLTLRKALFTIAVIWALALLIMCPS 176

Qy        180 AAVM----ECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAY 235
              |  :    |    : :  ||:     | | | :    |:| :  |   || || |: : |
Db        177 AVTLTVTREEHHFMLDARNRSYPLYSCWEAWPEKGMRKVYTAVLFAHIYLVPLALIVVMY 236

Qy        236 FQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRK 295
              :| |||   | ||                  | |:           |     :  | | :
Db        237 VRIARKLC--QAPGPAR--------------DTEE----------AVAEGGRTSRRRAR 269

Qy        296 TAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDRE----AVYACFTFSHWLVYANS 351
                ||::| | | | :||: || :|      : : |: :    :||| |  :||| : :|
Db        270 VVHMLVMVALFFTLSWLPLWVLLLLID----YGELSELQLHLLSVYA-FPLAHWLAFFHS 324

Qy        352 AANPIIYNFLSGKFREQFKAAF--SCCLP 378
              :|||||| : :  ||  |:|||    | |
Db        325 SANPIIYGYFNENFRRGFQAAFRAQLCWP 353


RESULT 10
NFF2_HUMAN
ID   NFF2_HUMAN     STANDARD;      PRT;   522 AA.
AC   Q9Y5X5; Q96RV1; Q9NR49;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor)
DE   (G-protein-coupled receptor HLWAR77).
GN   GPR74 OR NPGPR OR NPFF2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC   TISSUE=Fetal;
```

RX MEDLINE=99180505; PubMed=10079187;
RA Cikos S., Gregor P., Koppel J.;
RT "Sequence and tissue distribution of a novel G-protein-coupled
RT receptor expressed prominently in human placenta.";
RL Biochem. Biophys. Res. Commun. 256:352-356(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20408933; PubMed=10851242;
RA Elshourbagy N.A., Ames R.S., Fitzgerald L.R., Foley J.J.,
RA Chambers J.K., Szekeres P.G., Evans N.A., Schmidt D.B., Buckley P.T.,
RA Dytko G.M., Murdock P.R., Milligan G., Groarke D.A., Tan K.B.,
RA Shabon U., Nuthulaganti P., Wang D.Y., Wilson S., Bergsma D.J.,
RA Sarau H.M.;
RT "Receptor for the pain modulatory neuropeptides FF and AF is an orphan
RT G protein-coupled receptor.";
RL J. Biol. Chem. 275:25965-25971(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC TISSUE=Spinal cord;
RX MEDLINE=20564301; PubMed=11024015;
RA Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
RA Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhlani P.P.,
RA Raddatz R., Yao W.-J., Ogozalek K.L., Boyle N., Kouranova E.V.,
RA Quan Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
RA Borowsky B.;
RT "Identification and characterization of two G protein-coupled
RT receptors for neuropeptide FF.";
RL J. Biol. Chem. 275:39324-39331(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=20299143; PubMed=10837915;
RA Parker R.M.C., Copeland N.G., Eyre H.J., Liu M., Gilbert D.J.,
RA Crawford J., Couzens M., Sutherland G.R., Jenkins N.A., Herzog H.;
RT "Molecular cloning and characterisation of GPR74 a novel G-protein
RT coupled receptor closest related to the Y-receptor family.";
RL Brain Res. Mol. Brain Res. 77:199-208(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA Wang R., Evans J., Gould R., Austin C.P.;
RT "Identification and characterization of two cognate receptors for
RT mammalian FMRFamide-like neuropeptides.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RA Laemmle B.S., Schindler M., Beilmann M., Hamilton B.S., Doods H.N.,
RA Wieland H.A.;
RT "Cloning and characterization of the NPGP receptor and identification
RT of a novel short mRNA isoform in human hypothalamus.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Receptor for NPAF (A-18-F-amide) and NPFF (F-8-F-amide)
CC neuropeptides, also known as morphine-modulating peptides. Can
CC also be activated by a variety of naturally ocurring or synthetic
CC FMRF-amide like ligands. This receptor mediates its action by
CC association with G proteins that activate a phosphatidylinositol-

```
CC       calcium second messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1; Synonyms=long form;
CC         IsoId=Q9Y5X5-1; Sequence=Displayed;
CC       Name=2; Synonyms=short form;
CC         IsoId=Q9Y5X5-2; Sequence=VSP_001907;
CC       Name=3;
CC         IsoId=Q9Y5X5-3; Sequence=VSP_001908, VSP_001909;
CC       Name=4;
CC         IsoId=Q9Y5X5-4; Sequence=VSP_001910, VSP_001911;
CC   -!- TISSUE SPECIFICITY: Isoform 1 is abundant in placenta. Relatively
CC       highly expressed in thymus, testis, and small intestine. Expressed
CC       at low levels in several tissues including spleen, prostate,
CC       brain, heart, ovary, colon, kidney, lung, liver and pancreas and
CC       not expressed in skeletal muscle and leukocytes. Highest but
CC       relatively low level of isoform 2 in placenta and very low level
CC       in numerous tissues including adipose tissue and many brain
CC       regions. Isoform 3 is expressed in brain and heart and, at lower
CC       levels, in kidney, liver, lung and pancreas.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC       HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.
CC   -!- CAUTION: Ref.4 sequence differs from that shown due to a
CC       frameshift in position 503.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF119815; AAD22047.1; -.
DR   EMBL; AF257210; AAF87078.1; -.
DR   EMBL; AF268899; AAG41398.1; -.
DR   EMBL; AF236083; AAK58513.1; ALT_FRAME.
DR   EMBL; AF330053; AAK94197.1; -.
DR   EMBL; AJ311393; CAC85427.1; -.
DR   Genew; HGNC:4525; GPR74.
DR   MIM; 607449; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0008151; P:cell growth and/or maintenance; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0009582; P:perception of abiotic stimulus; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005397; NPFF_receptor2.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01572; NPFFRECEPTR2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
```

```
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN        1    147       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    148    168       1 (POTENTIAL).
FT   DOMAIN      169    184       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    185    205       2 (POTENTIAL).
FT   DOMAIN      206    221       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    222    242       3 (POTENTIAL).
FT   DOMAIN      243    262       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    263    283       4 (POTENTIAL).
FT   DOMAIN      284    319       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    320    340       5 (POTENTIAL).
FT   DOMAIN      341    377       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    378    398       6 (POTENTIAL).
FT   DOMAIN      399    413       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    414    434       7 (POTENTIAL).
FT   DOMAIN      435    522       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    220    308       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    300    300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1    102       Missing (in isoform 2).
FT                                /FTId=VSP_001907.
FT   VARSPLIC      1     99       Missing (in isoform 3).
FT                                /FTId=VSP_001908.
FT   VARSPLIC    100    100       R -> M (in isoform 3).
FT                                /FTId=VSP_001909.
FT   VARSPLIC    101    132       FIMNEKWDTNSSENWHPIWNVNDTKHHLYSDI -> MAIWK
FT                                HDVQDQWIGPGNICRSFSLYVSCNCCR (in isoform
FT                                4).
FT                                /FTId=VSP_001910.
FT   VARSPLIC    133    522       Missing (in isoform 4).
FT                                /FTId=VSP_001911.
FT   CONFLICT    466    466       A -> T (IN REF. 1 AND 4).
SQ   SEQUENCE   522 AA;  60270 MW;  40CB9FCD42F77041 CRC64;

  Query Match             22.5%;  Score 499.5;  DB 1;  Length 522;
  Best Local Similarity   33.4%;  Pred. No. 4.1e-27;
  Matches  120;  Conservative   59;  Mismatches  131;  Indels   49;  Gaps    7;

Qy         27 YEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFIV 86
              | |  : |:   ||:  |   : | :|  :| : ::|||:||  | || || |||| ||:
Db        129 YSDINITYV-NYYLHQPQVAAIFIISYFLIFFLCMMGNTVVCFIVMRNKHMHTVTNLFIL 187

Qy         87 NLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWY 146
              ||:::|:||  |:| :|| :|   | ||: :||:   :| :||: :| ||  ||:||:
Db        188 NLAISDLLVGIFCMPITLLDNIIAGWPFGNTMCKISGLVQGISVAASVFTLVAIAVDRFQ 247

Qy        147 AICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSV----- 201
               : :|   | | : |   |: || ::: || | | ::    |  |   | || |
Db        248 CVVYPFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH---VQEEKYYRVRLNSQNKTSP 304

Qy        202 ---CDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNW 258
                | | | :    ||| :  |   ||||| |: : | :|   |:   :| |       |
Db        305 VYWCREDWPNQEMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLFRAAVPHTGRKNQEQW 364
```

```
Qy        259 KRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLN 318
                |                                 | ::|  |||::| |:| | :||:  |
Db        365 HVVS----------------------------RKKQKIIKMLLIVALLFILSWLPLWTLM 396

Qy        319 VLKRVFGMFRQASDREA----VYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAF 373
              :|      :   |  |     :|  : |:||| : ||: ||||| | :  ||  |: ||
Db        397 MLSD----YADLSPNELQIINIY-IYPFAHWLAFGNSSVNPIIYGFFNENFRRGFQEAF 450


RESULT 11
NFF2_RAT
ID   NFF2_RAT       STANDARD;      PRT;   417 AA.
AC   Q9EQD2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor).
GN   NPGPR OR NPFF2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=20564301; PubMed=11024015;
RA   Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
RA   Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhlani P.P.,
RA   Raddatz R., Yao W.-J., Ogozalek K.L., Boyle N., Kouranova E.V.,
RA   Quan Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
RA   Borowsky B.;
RT   "Identification and characterization of two G protein-coupled
RT   receptors for neuropeptide FF.";
RL   J. Biol. Chem. 275:39324-39331(2000).
CC   -!- FUNCTION: Receptor for NPAF (A-18-F-amide) and NPFF (F-8-F-amide)
CC       neuropeptides, also known as morphine-modulating peptides. Can
CC       also be activated by a variety of naturally ocurring or synthetic
CC       FMRF-amide like ligands. This receptor mediates its action by
CC       association with G proteins that activate a phosphatidylinositol-
CC       calcium second messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC       HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF268900; AAG41399.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR005395; NPFF_receptor.
DR   InterPro; IPR005397; NPFF_receptor2.
```

```
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PRINTS; PR01572; NPFFRECEPTR2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     45       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     46     66       1 (POTENTIAL).
FT   DOMAIN       67     82       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     83    103       2 (POTENTIAL).
FT   DOMAIN      104    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    120    140       3 (POTENTIAL).
FT   DOMAIN      141    160       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    161    181       4 (POTENTIAL).
FT   DOMAIN      182    217       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    218    238       5 (POTENTIAL).
FT   DOMAIN      239    274       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    275    295       6 (POTENTIAL).
FT   DOMAIN      296    310       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    311    331       7 (POTENTIAL).
FT   DOMAIN      332    417       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    118    206       BY SIMILARITY.
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     20     20       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     31     31       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    198    198       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   417 AA;  47710 MW;  DFEB6AC63AF2AAD6 CRC64;

  Query Match             22.5%;  Score 499;  DB 1;  Length 417;
  Best Local Similarity   31.4%;  Pred. No. 3.6e-27;
  Matches  118;  Conservative   61;  Mismatches  135;  Indels   62;  Gaps    6;

Qy         25 PDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYF 84
              | | |  : |:   ||:      | |::|  :| : :||||:||  | || :| ||||:|
Db         25 PWYSDINITYM-NYYLHQPHVTAVFISSYFLIFFLCMVGNTVVCFVVIRNRYMHTVTNFF 83

Qy         85 IVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDR 144
              | ||:::|:||   |:| :|| :|   | || ::||:   :| :||: :| ||  ||:||
Db         84 IFNLAISDLLVGIFCMPITLLDNIIAGWPFGSSMCKISGLVQGISVAASVFTLVAIAVDR 143

Qy        145 WYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFS---- 200
              :  : :|   | | : |   |: || ::: || | | ::    |  |   | || |
Db        144 FRCVVYPFKPKLTVKTAFVMIVIIWGLAITIMTPSAIMLH---VQEEKYYRVRLSSHNKT 200

Qy        201 ----VCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVR 256
                 | | | :    :|| :  |   ||||| |: : | :|   |:      |    :
Db        201 STVYWCREDWPNQEMRRIYTTVLFATIYLAPLSLIVIMYARIGASLFKTSAHSTGKQRLE 260

Qy        257 NWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISV 316
              |                                  : ::|  |||: | |:| | :||:
Db        261 QW----------------------------HVSKKKQKVIKMLLTVALLFILSWLPLWT 291

Qy        317 LNVLK----------RVFGMFRQASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFR 366
              | :|           ||  ::            : |:||| : ||: ||||| | :  ||
Db        292 LMMLSDYADLSPNKLRVINIY----------VYPFAHWLAFCNSSVNPIIYGFFNENFR 340
```

```
Qy          367 EQFKAAFSCCLPGLGP 382
                |: ||  |   : |
Db          341 SGFQDAFQFCQKKVKP 356


RESULT 12
CCKR_RABIT
ID   CCKR_RABIT     STANDARD;      PRT;   427 AA.
AC   O97772;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN   CCKAR.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95002144; PubMed=7918628;
RA   Reuben M., Rising L., Prinz C., Hersey S., Sachs G.;
RT   "Cloning and expression of the rabbit gastric CCK-A receptor.";
RL   Biochim. Biophys. Acta 1219:321-327(1994).
CC   -!- FUNCTION: Receptor for cholecystokinin. Has a 1000 fold affinity
CC       for CCK rather than for gastrin. It modulates feeding and
CC       dopamine-induced behavior in the central and peripheral nervous
CC       system. This receptor mediates its action by association with G
CC       proteins that activate a phosphatidylinositol-calcium second
CC       messenger system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U86601; AAD11547.1; -.
DR   HSSP; P02699; 1F88.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Lipoprotein; Palmitate.
FT   DOMAIN        1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     42     67       1 (POTENTIAL).
FT   DOMAIN       68     77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78    104       2 (POTENTIAL).
FT   DOMAIN      105    115       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    116    137       3 (POTENTIAL).
FT   DOMAIN      138    157       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     158     178        4 (POTENTIAL).
FT   DOMAIN       179     210        EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     211     234        5 (POTENTIAL).
FT   DOMAIN       235     312        CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     313     333        6 (POTENTIAL).
FT   DOMAIN       334     348        EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     349     372        7 (POTENTIAL).
FT   DOMAIN       373     427        CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      10      10        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      24      24        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     190     190        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     114     196        BY SIMILARITY.
FT   LIPID        386     386        S-palmitoyl cysteine (By similarity).
SQ   SEQUENCE   427 AA;  47386 MW;  089FD10E2B86DB25 CRC64;

  Query Match             21.0%;  Score 466.5;  DB 1;  Length 427;
  Best Local Similarity   29.6%;  Pred. No. 6e-25;
  Matches  128;  Conservative   86;  Mismatches  167;  Indels   51;  Gaps   13;

Qy          8 GAQMGVPPGSREPSPVPPDYEDEFLRYLWRDYLYPKQYEW---VLIAAYVAVFVVALVGN 64
              |   |:||      |     ::| |  | :    |   ||   | |  |  :|:::::||
Db          9 GNASGIPP------PCELGLDNETLFCLDQP---PPSKEWQPAVQILLYSLIFLLSVLGN 59

Qy         65 TLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALCKVIPY 124
              |||   : ||  |||||| |:::|:::|:::   |:| :|: :: : ::|| ||||   |
Db         60 TLVITVLIRNKRMRTVTNIFLLSLAISDLMLCLFCMPFNLIPNLLKDFIFGSALCKTTTY 119

Qy        125 LQAVSVSVAVLTLSFIALDRWYAICHPLLFK--STARRARGSILGIWAVSLAIMVPQAAV 182
              |   ||||: | |  |:|:|: ||| ||  :   |   |   |   | :| ||| |
Db        120 LMGTSVSVSTLNLVAISLERYGAICKPLQSRVWQTKSHALKVIAATWCLSFAIMTPYPIY 179

Qy        183 MECSSVLPELANRTRLFSVCDERWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKL 242
                 |:::|      :  ::|      |:  : :|:   :: :| |  :| :||  |  :|
Db        180 ---SNLVPFTKTNNQTANMCRFLLPSDVMQQAWHTFLLLILFLIPGIVMMVAYGMISLEL 236

Qy        243 W-GRQIPGTTSALVRNWK-------------------RPSDQLGDLEQ--GLSGEPQPRA 280
              : | :   :     :  |                    :|: || :|:|  |  |    |
Db        237 YQGIKFDASQKKSAKERKASTGSGRFEDNDGCYLQRSKPTRQL-ELQQLSGGGGGRVSRI 295

Qy        281 RAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACF 340
              |:  :    : |:::  :||||:::::| ||::||   |  :    :   :::|
Db        296 RS-SSSAATLMAKKRVIRMLMVIVVLFFLCWMPIFSANAWRAYDTV---SAERRLSGTPI 351

Qy        341 TFSHWLVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSL 400
              :|   | | :|  ||||| |:: :||  | | | || |  ||      | ||: |  :
Db        352 SFILLLSYTSSCVNPIIYCFMNRRFRLGFMATFPCC-PNPGP-------PGPRAEAGEEEE 404

Qy        401 SLQSRCSVSKIS 412
                :| |:|: |
Db        405 GRTTRASLSRYS 416


RESULT 13
NY2R_PIG
ID   NY2R_PIG       STANDARD;      PRT;   382 AA.
AC   O02836; Q9TSI1;
DT   15-JUL-1998 (Rel. 36, Created)
```

```
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN   NPY2R.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=99017380; PubMed=9802394;
RA   Malmstroem R.E., Hoekfelt T., Bjoerkman J.-A., Nihlen C., Bystroem M.,
RA   Ekstrand A.J., Lundberg J.M.;
RT   "Characterization and molecular cloning of vascular neuropeptide Y
RT   receptor subtypes in pig and dog.";
RL   Regul. Pept. 75:55-70(1998).
RN   [2]
RP   REVISION TO 207.
RA   Ekstrand A.J.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21184969; PubMed=11287083;
RA   Larhammar D., Wraith A., Berglund M.M., Holmberg S.K., Lundell I.;
RT   "Origins of the many NPY-family receptors in mammals.";
RL   Peptides 22:295-307(2001).
CC   -!- FUNCTION: Receptor for neuropeptide Y and peptide YY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC       Highest to tachykinins receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF005780; AAC26670.2; -.
DR   EMBL; AF106082; AAD13777.1; ALT_INIT.
DR   HSSP; P02699; 1F88.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     50       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     51     73       1 (POTENTIAL).
FT   DOMAIN       74     83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     84    105       2 (POTENTIAL).
FT   DOMAIN      106    125       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    126    147       3 (POTENTIAL).
FT   DOMAIN      148    167       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     168     188        4 (POTENTIAL).
FT   DOMAIN       189     215        EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     216     241        5 (POTENTIAL).
FT   DOMAIN       242     269        CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     270     292        6 (POTENTIAL).
FT   DOMAIN       293     305        EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     306     329        7 (POTENTIAL).
FT   DOMAIN       330     382        CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      11      11        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     124     204        BY SIMILARITY.
FT   LIPID        343     343        S-palmitoyl cysteine (Potential).
FT   CONFLICT       4       4        I -> L (IN REF. 3).
FT   CONFLICT     179     179        L -> V (IN REF. 3).
FT   CONFLICT     215     215        I -> N (IN REF. 3).
FT   CONFLICT     366     366        A -> V (IN REF. 3).
SQ   SEQUENCE   382 AA;  42556 MW;  5C01FAF3A0423858 CRC64;

  Query Match             20.3%;  Score 451;  DB 1;  Length 382;
  Best Local Similarity   28.6%;  Pred. No. 6.1e-24;
  Matches  118;  Conservative   75;  Mismatches  141;  Indels   78;  Gaps   12;

Qy          1 MEPSATPGAQMGVPPGSREP-SPVPPDYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVV 59
              ||||          ||    |   : || | |      :|     : : :|| || :: ::
Db         19 MEPSG---------PGHTTPRGELAPDSEPEL-----KDSTKLIEVQIILILAYCSIILL 64

Qy         60 ALVGNTLVCLAVWRNHHMRTVTNYFIVNLSLADVLVTAICLPASLLVDITESWLFGHALC 119
               :|||:||   | :   ||||||:|| ||::||:||  :||| :|   :   |  |  ||
Db         65 GVVGNSLVIHVVIKFKSMRTVTNFFIANLAVADLLVNTLCLPFTLTYTLMGEWKMGPVLC 124

Qy        120 KVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSILGI-WAVSLAIMVP 178
               ::|| | ::| |: :||: |||||   | : |  |  ::|    |:|: | :|  :  |
Db        125 HLVPYAQGLAVQVSTITLTVIALDRHRCIVYHLESK-ISKRISFLIIGLAWGISALLASP 183

Qy        179 QAAVMECS--SVLPELANRTRLFSVCDERWADD---LYPKIYHSCFFIVTYLAPLGLMAM 233
               |   | |   ::|:          | |:|  :   :|  :|     :: |: |||:::
Db        184 LAIFREYSLIEIIPDFE-----IVACTEKWPGEEKSIYGTVYSLSSLLILYVLPLGIISF 238

Qy        234 AYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLSGEPQPRARAFLAEVKQMRAR 293
               :| :|: ||     ||                     |::     |             |
Db        239 SYARIWSKLKNHVSPG--------------------GVNDHYHQR-------------R 264

Qy        294 RKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDREAVYACFTFSHWLVYANSAA 353
               :|| |||: |::|||: :||   |:  :    : .|  | :     ||  | :   :: |
Db        265 QKTTKMLVCVVVVFAVSWLP---LHAFQLAVDIDSQVLDLKEYKLIFTVFHIIAMCSTFA 321

Qy        354 NPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLKAPSPRSSASHKSLSLQSR 405
               ||::| :::  :|: | :|| |                |  | |  :|: |:
Db        322 NPLLYGWMNSNYRKAFLSAFRC-------------EQRLDAIHSEVSMTSK 359


RESULT 14
NYR_DROME
ID   NYR_DROME      STANDARD;      PRT;   449 AA.
AC   P25931;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
```

```
DE   Neuropeptide Y receptor (NPY-R) (PR4 receptor).
GN   NEPYR.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92112730; PubMed=1370455;
RA   Li X.-J., Wu Y.-N., North R.A., Forte M.;
RT   "Cloning, functional expression, and developmental regulation of a
RT   neuropeptide Y receptor from Drosophila melanogaster.";
RL   J. Biol. Chem. 267:9-12(1992).
CC   -!- FUNCTION: Receptor for neuropeptide Y and peptide YY.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- DEVELOPMENTAL STAGE: At low levels during early embryonic stages,
CC       its expression increases later and reaches the highest level
CC       during late stages of embryogenesis. Subsequently, PR4 levels are
CC       reduced during larval stages and increase during pupal stages.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC       Highest to tachykinins receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M81490; AAA28727.1; -.
DR   PIR; A41738; A41738.
DR   FlyBase; FBgn0004842; NepYr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     85       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     86    111       1 (POTENTIAL).
FT   DOMAIN      112    120       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    121    145       2 (POTENTIAL).
FT   DOMAIN      146    165       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    166    187       3 (POTENTIAL).
FT   DOMAIN      188    203       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    204    232       4 (POTENTIAL).
FT   DOMAIN      233    256       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    257    282       5 (POTENTIAL).
FT   DOMAIN      283    309       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    310    330       6 (POTENTIAL).
FT   DOMAIN      331    347       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    348    372       7 (POTENTIAL).
FT   DOMAIN      373    449       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     34     34       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD     70      70       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   449 AA;  51885 MW;  1ADB32EFD50146C3 CRC64;

  Query Match             20.2%;  Score 447.5;  DB 1;  Length 449;
  Best Local Similarity   29.4%;  Pred. No. 1.2e-23;
  Matches  109;  Conservative   63;  Mismatches  138;  Indels   61;  Gaps    9;

Qy         26 DYEDEFLRYLWRDYLYPKQYEWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRTVTNYFI 85
              ||:      |  |      :: ::   |: :|: ||:|| || |:    |||||||||
Db         75 DYD-----LLSEDMWSSAYFKIIVYMLYIPIFIFALIGNGTVCYIVYSTPRMRTVTNYFI 129

Qy         86 VNLSLADVLVTAICLPASLL-VDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDR 144
               :|:: |:|::  | |:| : : |   | || |||  : | ||||| |:  ||  |::||
Db        130 ASLAIGDILMSFFCEPSSFISLFILNYWPFGLALCHFVNYSQAVSVLVSAYTLVAISIDR 189

Qy        145 WYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELANRTRLFSVCDE 204
              : ||  ||  : | | |   | |:| ::||  :|   |      :    :   : :| |
Db        190 YIAIMWPLKPRITKRYATFIIAGVWFIALATALPIPIVSGLDIPMSPWHTKCEKY-ICRE 248

Qy        205 RWADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQ 264
               |       |    | : :: |||::   | :|  ::| :: ||         :   ||
Db        249 MWPSRSQEYYYTLSLFALQFVVPLGVLIFTYARITIRVWAKRPPGEA-------ETNRDQ 301

Qy        265 LGDLEQGLSGEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVF 324
                                 :  |::||  ||:: |::||  |:|| ::| :|
Db        302 ----------------------RMARSKRKMVKMMLTVVIVFTCCWLPFNILQLL---- 334

Qy        325 GMFRQASDREAV------YACFTFSHWLVYANSAANPIIYNFLSGKFREQFKAAFS---- 374
                  :| |        |  | | |||  ::   ||||| ::: :||  |
Db        335 -----LNDEEFAHWDPLPYVWFAF-HWLAMSHCCYNPIIYCYMNARFRSGFVQLMHRMPG 388

Qy        375 ----CCLPGLG 381
                  |||  :|
Db        389 LRRWCCLRSVG 399


RESULT 15
OX1R_PIG
ID   OX1R_PIG       STANDARD;      PRT;    86 AA.
AC   O97661;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 1 (Ox1r) (Hypocretin receptor type 1) (Fragment).
GN   HCRTR1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RA   Matteri R.L., Dyer C.J.;
RT   "Partial cDNA sequence of the porcine type 1 orexin receptor.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Moderately selective excitatory receptor for orexin-A
CC       and, with a lower affinity, for orexin-B neuropeptide. Seems to be
```

```
CC       exclusively coupled to the G(q) subclass of heteromeric G
CC       proteins, which activates the phospholipase C mediated signaling
CC       cascade (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF097995; AAC68703.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   NON_TER       1      1
FT   TRANSMEM     <1      8       3 (POTENTIAL).
FT   DOMAIN        9     30       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     31     51       4 (POTENTIAL).
FT   DOMAIN       52     82       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     83    >86       5 (POTENTIAL).
FT   NON_TER      86     86
SQ   SEQUENCE   86 AA;  9853 MW;  C34177B3BC8590E8 CRC64;

  Query Match             19.7%;  Score 437;  DB 1;  Length 86;
  Best Local Similarity   96.5%;  Pred. No. 1.4e-23;
  Matches   83;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        135 LTLSFIALDRWYAICHPLLFKSTARRARGSILGIWAVSLAIMVPQAAVMECSSVLPELAN 194
              ||||||||||||||| |||||||||||||||||||||| |:||||||||||||||||||
Db          1 LTLSFIALDRWYAIYHPLLFKSTARRARGSILGIWAVSPAVMVPQAAVMECSSVLPELAN 60

Qy        195 RTRLFSVCDERWADDLYPKIYHSCFF 220
              ||||||||||||||||||||||||||
Db         61 RTRLFSVCDERWADDLYPKIYHSCFF 86


Search completed: September 28, 2004, 09:50:25
Job time : 15 secs
```